

日 本 国 特 許 庁
JAPAN PATENT OFFICE

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office

出 願 年 月 日
Date of Application:

2002年12月 4日

出 願 番 号
Application Number:

特願2002-352865

[ST.10/C]:

[JP2002-352865]

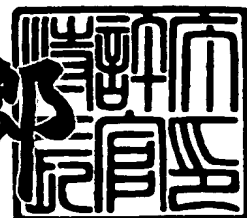
出 願 人
Applicant(s):

パイオニア株式会社

2003年 6月24日

特 許 庁 長 官
Commissioner,
Japan Patent Office

太田 信一郎



出証番号 出証特2003-3049530

【書類名】 特許願

【整理番号】 57P0355

【提出日】 平成14年12月 4日

【あて先】 特許庁長官殿

【国際特許分類】 G10L 15/10
G06F 17/30
G10H 1/38

【発明の名称】 楽曲構造検出装置及び方法

【請求項の数】 10

【発明者】
【住所又は居所】 埼玉県鶴ヶ島市富士見6丁目1番1号 パイオニア株式会社 総合研究所内

【氏名】 我山 真一

【特許出願人】
【識別番号】 000005016
【氏名又は名称】 パイオニア株式会社

【代理人】
【識別番号】 100079119
【弁理士】
【氏名又は名称】 藤村 元彦

【手数料の表示】
【予納台帳番号】 016469
【納付金額】 21,000円

【提出物件の目録】
【物件名】 明細書 1
【物件名】 図面 1
【物件名】 要約書 1
【包括委任状番号】 9006557

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 楽曲構造検出装置及び方法

【特許請求の範囲】

【請求項 1】 楽曲の和音の時系列変化を示す和音進行楽曲データに応じてその楽曲の構造を検出する楽曲構造検出装置であって、

前記和音進行楽曲データ中の各和音の位置から連続する所定数の和音からなる部分楽曲データを生成する部分楽曲データ生成手段と、

前記部分楽曲データ各々と前記和音進行楽曲データとを前記和音進行楽曲データ中の各和音の位置から和音変化時の和音の根音変化量と変化後の和音の属性とについて比較して前記複数の楽曲毎の類似度を算出する比較手段と、

前記部分楽曲データ毎に前記比較手段によって算出された類似度各々に応じて類似度が所定値より高いピーク値となった前記和音進行楽曲データ中の和音の位置を検出する和音位置検出手段と、

前記和音進行楽曲データ中の和音の位置毎に前記部分楽曲データ全てについて前記類似度が前記所定値より高いピーク値となった回数を算出し、その和音の位置毎の算出回数に応じて楽曲構造を示す検出出力を生成する出力手段と、を備えたことを特徴とする楽曲構造検出装置。

【請求項 2】 前記比較手段は、前記部分楽曲データ各々と前記和音進行楽曲データとを前記和音進行楽曲データ中の各和音の位置から和音変化時の和音の根音変化量と変化後の和音の属性とに加えて和音変化前後の和音の時間的長さの比について比較して前記複数の楽曲毎の類似度を算出することを特徴とする請求項 1 記載の楽曲構造検出装置。

【請求項 3】 前記比較手段は、前記部分楽曲データ各々と前記和音進行楽曲データとを時間的に前後に飛び越して比較することを特徴とする請求項 1 記載の楽曲構造検出装置。

【請求項 4】 前記比較手段は、前記部分楽曲データ各々が示す和音変化後の和音と前記和音進行楽曲データが示す和音変化後の和音とが関係調にあるときその双方の和音変化後の和音を同一の和音とみなすことを特徴とする請求項 1 記載の楽曲構造検出装置。

【請求項 5】 前記部分楽曲データ各々と前記和音進行楽曲データとは和音変化時点毎に 2 つの和音を第 1 及び第 2 和音候補として有し、

前記比較手段は、前記部分楽曲データ各々の第 1 及び第 2 和音候補と前記和音進行楽曲データの第 1 及び第 2 和音候補とを相互に比較することを特徴とする請求項 1 記載の楽曲構造検出装置。

【請求項 6】 楽曲を示す入力オーディオ信号を所定の時間毎に周波数成分の大きさを示す周波数信号に変換する周波数変換手段と、

前記周波数変換手段によって得られた周波数信号から平均律の各音に対応した周波数成分を前記所定の時間毎に抽出する成分抽出手段と、

前記成分抽出手段によって抽出された各音に対応した周波数成分のうちのレベル合計が大となる 3 つの周波数成分の組によって各々形成される 2 つの和音を前記第 1 及び第 2 和音候補として検出する和音候補検出手段と、

前記和音候補検出手段によって繰り返し検出された前記第 1 及び第 2 和音候補各々の列を平滑化処理して前記蓄積手段に記憶させるべき前記和音進行楽曲データを生成する平滑化手段と、を備えたことを特徴とする請求項 5 記載の楽曲構造検出装置。

【請求項 7】 前記比較手段は、前記和音進行楽曲データの終了部分に前記所定数の仮想和音だけの仮想データを付加して前記部分楽曲データ各々との比較に用いることを特徴とする請求項 1 記載の楽曲構造検出装置。

【請求項 8】 前記出力手段は、前記和音進行楽曲データ中の和音の位置毎の算出回数が最高回数の部分の楽曲音を再生して出力する請求項 1 記載の楽曲構造検出装置。

【請求項 9】 楽曲の和音の時系列変化を示す和音進行楽曲データに応じてその楽曲の構造を検出する楽曲構造検出方法であって、

前記和音進行楽曲データ中の各和音の位置から連続する所定数の和音からなる部分楽曲データを生成する部分楽曲データ生成ステップと、

前記部分楽曲データ各々と前記和音進行楽曲データとを前記和音進行楽曲データ中の各和音の位置から和音変化時の和音の根音変化量と変化後の和音の属性とについて比較して前記複数の楽曲毎の類似度を算出する比較ステップと、

前記部分楽曲データ毎に前記比較ステップにおいて算出された類似度各々に応じて類似度が所定値より高いピーク値となった前記和音進行楽曲データ中の和音の位置を検出する和音位置検出ステップと、

前記和音進行楽曲データ中の和音の位置毎に前記部分楽曲データ全てについて前記類似度が前記所定値より高いピーク値となった回数を算出し、その和音の位置毎の算出回数に応じて楽曲構造を示す検出出力を生成する出力ステップと、を備えたことを特徴とする楽曲構造検出方法。

【請求項 1 0】 楽曲の和音の時系列変化を示す和音進行楽曲データに応じてその楽曲の構造を検出する方法を実行するコンピュータ読取可能なプログラムであって、

前記和音進行楽曲データ中の各和音の位置から連続する所定数の和音からなる部分楽曲データを生成する部分楽曲データ生成ステップと、

前記部分楽曲データ各々と前記和音進行楽曲データとを前記和音進行楽曲データ中の各和音の位置から和音変化時の和音の根音変化量と変化後の和音の属性とについて比較して前記複数の楽曲毎の類似度を算出する比較ステップと、

前記部分楽曲データ毎に前記比較ステップにおいて算出された類似度各々に応じて類似度が所定値より高いピーク値となった前記和音進行楽曲データ中の和音の位置を検出する和音位置検出ステップと、

前記和音進行楽曲データ中の和音の位置毎に前記部分楽曲データ全てについて前記類似度が前記所定値より高いピーク値となった回数を算出し、その和音の位置毎の算出回数に応じて楽曲構造を示す検出出力を生成する出力ステップと、を備えたことを特徴とするプログラム。

【発明の詳細な説明】

【0 0 0 1】

【発明が属する技術分野】

本発明は、楽曲の和音の時系列変化を示すデータに応じてその楽曲の構造を検出する楽曲構造検出装置及び方法に関する。

【0 0 0 2】

【従来の技術】

ポピュラー音楽の楽曲においては、フレーズ（楽句）がイントロ、Aメロ、Bメロ、サビのように表現され、Aメロ、Bメロやサビの各フレーズは楽曲中で通常何回か繰り返される。楽曲中のいわゆる盛り上がり部分であるサビのフレーズは、ラジオやテレビの音楽番組やコマーシャルで最も演奏されるところである。このようなフレーズは放送する際にその楽曲音を実際に聴取して判断することが一般的である。

【 0 0 0 3 】

【発明が解決しようとする課題】

ところで、楽曲のサビ等のフレーズがどのように繰り返されているかなどの楽曲全体の構造を知ることができれば、サビ部分に限らず、他の繰り返しフレーズ部分を容易に選択的に演奏することができる。しかしながら、従来、楽曲全体の構造を自動的に検出する装置はなく、上記したように利用者が実際に聴取して判断するしかなかった。

【 0 0 0 4 】

そこで、本発明が解決しようとする課題には、上記の問題点が一例として挙げられ、繰り返し部分を含む楽曲の構造を簡単な構成で適切に検出することができる楽曲構造検出装置及び方法を提供することが本発明の目的である。

【 0 0 0 5 】

【課題を解決するための手段】

本発明の楽曲構造検出装置は、楽曲の和音の時系列変化を示す和音進行楽曲データに応じてその楽曲の構造を検出する楽曲構造検出装置であって、前記和音進行楽曲データ中の各和音の位置から連続する所定数の和音からなる部分楽曲データを生成する部分楽曲データ生成手段と、前記部分楽曲データ各々と前記和音進行楽曲データとを前記和音進行楽曲データ中の各和音の位置から和音変化時の和音の根音変化量と変化後の和音の属性とについて比較して前記複数の楽曲毎の類似度を算出する比較手段と、前記部分楽曲データ毎に前記比較手段によって算出された類似度各々に応じて類似度が所定値より高いピーク値となった前記和音進行楽曲データ中の和音の位置を検出する和音位置検出手段と、前記和音進行楽曲データ中の和音の位置毎に前記部分楽曲データ全てについて前記類似度が前記所

定値より高いピーク値となった回数を算出し、その和音の位置毎の算出回数に応じて楽曲構造を示す検出出力を生成する出力手段と、を備えたことを特徴としている。

【0006】

本発明の楽曲構造検出装置は、楽曲の和音の時系列変化を示す和音進行楽曲データに応じてその楽曲の構造を検出する楽曲構造検出方法であって、前記和音進行楽曲データ中の各和音の位置から連続する所定数の和音からなる部分楽曲データを生成する部分楽曲データ生成ステップと、前記部分楽曲データ各々と前記和音進行楽曲データとを前記和音進行楽曲データ中の各和音の位置から和音変化時の和音の根音変化量と変化後の和音の属性とについて比較して前記複数の楽曲毎の類似度を算出する比較ステップと、前記部分楽曲データ毎に前記比較ステップにおいて算出された類似度各々に応じて類似度が所定値より高いピーク値となった前記和音進行楽曲データ中の和音の位置を検出する和音位置検出ステップと、前記和音進行楽曲データ中の和音の位置毎に前記部分楽曲データ全てについて前記類似度が前記所定値より高いピーク値となった回数を算出し、その和音の位置毎の算出回数に応じて楽曲構造を示す検出出力を生成する出力ステップと、を備えたことを特徴としている。

【0007】

本発明のプログラムは、楽曲の和音の時系列変化を示す和音進行楽曲データに応じてその楽曲の構造を検出する方法を実行するコンピュータ読取可能なプログラムであって、前記和音進行楽曲データ中の各和音の位置から連続する所定数の和音からなる部分楽曲データを生成する部分楽曲データ生成ステップと、前記部分楽曲データ各々と前記和音進行楽曲データとを前記和音進行楽曲データ中の各和音の位置から和音変化時の和音の根音変化量と変化後の和音の属性とについて比較して前記複数の楽曲毎の類似度を算出する比較ステップと、前記部分楽曲データ毎に前記比較ステップにおいて算出された類似度各々に応じて類似度が所定値より高いピーク値となった前記和音進行楽曲データ中の和音の位置を検出する和音位置検出ステップと、前記和音進行楽曲データ中の和音の位置毎に前記部分楽曲データ全てについて前記類似度が前記所定値より高いピーク値となった回数

を算出し、その和音の位置毎の算出回数に応じて楽曲構造を示す検出出力を生成する出力ステップと、を備えたことを特徴としている。

【0008】

【発明の実施の形態】

以下、本発明の実施例を図面を参照しつつ詳細に説明する。

図1は本発明を適用した楽曲処理システムを示している。この楽曲処理システムは、楽曲入力装置1、操作入力装置2、和音解析装置3、データ蓄積装置4、5、一時記憶メモリ6、和音進行比較装置7、繰り返し構造検出装置8、表示装置9、楽曲再生装置10、デジタル／アナログ変換装置11及びスピーカ12を備えている。

【0009】

楽曲入力装置1は和音解析装置3及びデータ蓄積装置5に接続され、デジタル化されたオーディオ信号（例えば、PCMデータ）を再生する装置であり、例えば、CDプレーヤである。操作入力装置2は本システムに対してユーザが操作してデータや指令を入力するための装置である。操作入力装置2の出力は和音解析装置3、和音進行比較装置7、繰り返し構造検出装置8及び楽曲再生装置10に接続されている。データ蓄積装置4には楽曲入力装置1から供給された楽曲データ（PCMデータ）がファイルとして記憶される。

【0010】

和音解析装置3は、供給された楽曲データの和音を後述する和音解析動作によって解析する。一時記憶メモリ6には和音解析装置3によって解析された楽曲データの各和音が第1及び第2和音候補として一時的に記憶される。データ蓄積装置5には和音解析装置3によって解析されて和音進行楽曲データが楽曲毎にファイルとして記憶される。

【0011】

和音進行比較装置7は、データ蓄積装置5に記憶された和音進行楽曲データとその和音進行楽曲データ中の一部分である部分楽曲データとを後述するように比較して類似度を算出する。繰り返し構造検出装置8は和音進行比較装置7の比較結果を用いて楽曲の繰り返し部分を検出する。

表示装置 9 には繰り返し構造検出装置 8 によって検出された繰り返し部分を含む楽曲構造が表示される。

【 0 0 1 2 】

楽曲再生装置 1 0 は、繰り返し構造検出装置 8 によって検出された繰り返し部分の楽曲データをデータ蓄積装置 4 から読み出して再生し、デジタルオーディオ信号として順次出力する。デジタル／アナログ変換装置 1 1 は楽曲再生装置 1 0 によって再生されたデジタルオーディオ信号をアナログオーディオ信号に変換してスピーカ 1 2 に供給する。

【 0 0 1 3 】

和音解析装置 3、和音進行比較装置 7、繰り返し構造検出装置 8 及び楽曲再生装置 1 0 各々は操作入力装置 2 からの指令に応じて動作する。

次に、かかる構成の楽曲処理システムの動作について説明する。

ここでは楽曲入力装置 1 から出力される楽曲音を示すデジタルオーディオ信号が和音解析装置 3 に供給されたとする。

【 0 0 1 4 】

上記した和音解析動作としては前処理、本処理及び後処理がある。和音解析装置 3 は前処理として周波数誤差検出動作を行う。

周波数誤差検出動作においては、図 2 に示すように、時間変数 T 及び帯域データ $F(N)$ が 0 に初期化され、更に変数 N の範囲が $-3 \sim 3$ の如く初期設定される（ステップ S 1）。入力デジタル信号に対してフーリエ変換によって周波数変換を 0.2 秒間隔で行うことによって周波数情報 $f(T)$ が得られる（ステップ S 2）。

【 0 0 1 5 】

今回の $f(T)$ 、前回の $f(T-1)$ 及び前々回の $f(T-2)$ を用いて移動平均処理が行われる（ステップ S 3）。この移動平均処理では、0.6 秒以内では和音が変わることが少ないという仮定で過去 2 回分の周波数情報が用いられる。移動平均処理は次式によって演算される。

$$f(T) = (f(T) + f(T-1) / 2.0 + f(T-2) / 3.0) / 3.0 \quad \dots\dots (1)$$

ステップ S 3 の実行後、変数 N が - 3 に設定され（ステップ S 4）、その変数 N は 4 より小であるか否かが判別される（ステップ S 5）。 $N < 4$ の場合には、移動平均処理後の周波数情報 $f(T)$ から周波数成分 $f_1(T) \sim f_5(T)$ が各々抽出される（ステップ S 6 ~ S 1 0）。周波数成分 $f_1(T) \sim f_5(T)$ は、 $(110.0 + 2 \times N)$ Hz を基本周波数とした 5 オクターブ分の平均律の 1 2 音のものである。1 2 音は A, A #, B, C, C #, D, D #, E, F, F #, G, G # である。図 3 は A 音を 1.0 とした場合の 1 2 音及び 1 オクターブ高い A 音各々の周波数比を示している。ステップ S 6 の $f_1(T)$ は A 音を $(110.0 + 2 \times N)$ Hz とし、ステップ S 7 の $f_2(T)$ は A 音を $2 \times (110.0 + 2 \times N)$ Hz とし、ステップ S 8 の $f_3(T)$ は A 音を $4 \times (110.0 + 2 \times N)$ Hz とし、ステップ S 9 の $f_4(T)$ は A 音を $8 \times (110.0 + 2 \times N)$ Hz とし、ステップ S 1 0 の $f_5(T)$ は A 音を $16 \times (110.0 + 2 \times N)$ Hz としている。

【 0 0 1 6 】

ステップ S 6 ~ S 1 0 の実行後、周波数成分 $f_1(T) \sim f_5(T)$ は 1 オクターブ分の帯域データ $F'(T)$ に変換される（ステップ S 1 1）。帯域データ $F'(T)$ は、

$$F'(T) = f_1(T) \times 5 + f_2(T) \times 4 + f_3(T) \times 3 + f_4(T) \times 2 + f_5(T) \quad \dots\dots (2)$$

の如く表される。すなわち、周波数成分 $f_1(T) \sim f_5(T)$ 各々は個別に重み付けされた後、加算される。1 オクターブの帯域データ $F'(T)$ は、帯域データ $F(N)$ に加算される（ステップ S 1 2）。その後、変数 N には 1 が加算され（ステップ S 1 3）、そして、ステップ S 5 が再度実行される。

【 0 0 1 7 】

ステップ S 6 ~ S 1 3 の動作は、ステップ S 5 において N が 4 より小、すなわち - 3 ~ + 3 の範囲であると判断される限り繰り返される。これによって音成分 $F(N)$ は - 3 ~ + 3 の範囲の音程誤差を含む 1 オクターブ分の周波数成分となる。

ステップ S 5 において $N \geq 4$ と判別された場合には、変数 T が所定値 M より小であるか否かが判別される（ステップ S 1 4）。 $T < M$ の場合には、変数 T に 1

が加算され（ステップ S 1 5）、ステップ S 2 が再度実行される。M 回分の周波数変換による周波数情報 $f(T)$ に対して変数 N 毎の帯域データ $F(N)$ が算出される。

【0018】

ステップ S 1 4 において $T \geq M$ と判別された場合には、変数 N 毎の 1 オクターブ分の帯域データ $F(N)$ のうちの各周波数成分の総和が最大値となる $F(N)$ が検出され、その検出 $F(N)$ の N が誤差値 X として設定される（ステップ S 1 6）。

この前処理によって誤差値 X を求めることによってオーケストラの演奏音等の楽曲音全体の音程が平均律と一定の差をもっている場合に、それを補償して後述の和音解析の本処理を行うことができる。

【0019】

前処理の周波数誤差検出動作が終了すると、和音解析動作の本処理が行われる。なお、誤差値 X が既に分かっている場合やその誤差を無視できる場合には、前処理は省略しても良い。本処理では楽曲全部について和音解析が行われるために楽曲の最初の部分から入力デジタル信号は和音解析装置 3 に供給されるところ。

【0020】

本処理においては、図 4 に示すように、入力デジタル信号に対してフーリエ変換によって周波数変換を 0.2 秒間隔で行うことによって周波数情報 $f(T)$ が得られる（ステップ S 2 1）。このステップ S 2 1 が周波数変換手段に対応する。そして、今回の $f(T)$ 、前回の $f(T-1)$ 及び前々回の $f(T-2)$ を用いて移動平均処理が行われる（ステップ S 2 2）。ステップ S 2 1 及び S 2 2 は上記したステップ S 2 及び S 3 と同様に実行される。

【0021】

ステップ S 2 2 の実行後、移動平均処理後の周波数情報 $f(T)$ から周波数成分 $f_1(T) \sim f_5(T)$ が各々抽出される（ステップ S 2 3 ～ S 2 7）。上記したステップ S 6 ～ S 1 0 と同様に、周波数成分 $f_1(T) \sim f_5(T)$ は、 $(110.0 + 2 \times N) \text{Hz}$ を基本周波数とした 5 オクターブ分の平均律の 12 音 A, A#, B, C, C#, D, D#, E, F, F#, G, G# である。ステップ S 2 3 の $f_1(T)$

T)はA音を $(110.0 + 2 \times N)$ Hzとし、ステップS 2 4の $f_2(T)$ はA音を $2 \times (110.0 + 2 \times N)$ Hzとし、ステップS 2 5の $f_3(T)$ はA音を $4 \times (110.0 + 2 \times N)$ Hzとし、ステップS 2 6の $f_4(T)$ はA音を $8 \times (110.0 + 2 \times N)$ Hzとし、ステップS 2 7の $f_5(T)$ はA音を $16 \times (110.0 + 2 \times N)$ Hzとしている。ここで、NはステップS 1 6で設定されたXである。

【 0 0 2 2 】

ステップS 2 3～S 2 7の実行後、周波数成分 $f_1(T) \sim f_5(T)$ は1オクターブ分の帯域データ $F'(T)$ に変換される(ステップS 2 8)。このステップS 2 8も上記のステップS 1 1と同様に式(2)を用いて実行される。帯域データ $F'(T)$ は各音成分を含むことになる。ステップS 2 3～S 2 8が成分抽出手段に相当する。

【 0 0 2 3 】

ステップS 2 8の実行後、帯域データ $F'(T)$ 中の各音成分のうちの強度レベルが大きいものから6音が候補として選択され(ステップS 2 9)、その6音候補から2つの和音 M_1 、 M_2 が作成される(ステップS 3 0)。候補の6音のうちから1つの音を根音(ルート)として3音からなる和音を作成される。すなわち ${}_6C_3$ 通りの組み合わせの和音が考慮される。各和音を構成する3音のレベルが加算され、その加算結果の値が最大となった和音が第1和音候補 M_1 とされ、加算結果の値が2番目に大きい和音が第2和音候補 M_2 とされる。

【 0 0 2 4 】

帯域データ $F'(T)$ の各音成分が図5に示すように12音に対する強度レベルを示す場合には、ステップS 2 9ではA、E、C、G、B、Dの6音が選択される。その6音A、E、C、G、B、Dのうちの3音から作成される3和音は、(A、C、E)からなる和音 A_m 、(音C、E、G)からなる和音C、(音E、B、G)からなる和音 E_m 、(音G、B、D)からなる和音G、……の如くである。和音 A_m (音A、C、E)の合計強度レベルは12、和音C(音C、E、G)の合計強度レベルは9、和音 E_m (音E、B、G)の合計強度レベルは7、和音G(音G、B、D)の合計強度レベルは4である。よって、ステップS 3 0では和音 A_m の合計強度レベル12が最大となるので、第1和音候補 M_1 として和音 A_m が設定さ

れ、和音Cの合計強度レベル7が2番目に大きいので、第2和音候補M2として和音Cが設定される。

【0025】

また、帯域データF'(T)の各音成分が図6に示すように12音に対する強度レベルを示す場合には、ステップS29ではC, G, A, E, B, Dの6音が選択される。その6音C, G, A, E, B, Dのうちの3音から作成される3和音は、(音C, E, G)からなる和音C、(A, C, E)からなる和音A_m、(音E, B, G)からなる和音E_m、(音G, B, D)からなる和音G、……の如くである。和音C(音C, E, G)の合計強度レベルは11、和音A_m(音A, C, E)の合計強度レベルは10、和音E_m(音E, B, G)の合計強度レベルは7、和音G(音G, B, D)の合計強度レベルは6である。よって、ステップS30では和音Cの合計強度レベル11が最大となるので、第1和音候補M1として和音Cが設定され、和音A_mの合計強度レベル10が2番目に大きいので、第2和音候補M2として和音A_mが設定される。

【0026】

和音を構成する音は3音に限らず、セブンスやディミニッシュセブンス等の4音もある。4音からなる和音に対しては図7に示すように3音からなる2つ以上の和音に分類されるとしている。よって、4音からなる和音に対しても3音からなる和音と同様に、帯域データF'(T)の各音成分の強度レベルに応じて2つの和音候補を設定することができる。

【0027】

ステップS30の実行後、ステップS30において設定された和音候補数があるか否かが判別される(ステップS31)。ステップS30では少なくとも3つの音を選択するだけの強度レベルに差がない場合には和音候補が全く設定されないことになるので、ステップS31の判別が行われる。和音候補数>0である場合には、更に、その和音候補数が1より大であるか否かが判別される(ステップS32)。

【0028】

ステップS31において和音候補数=0と判別された場合には前回T-1(約

0.2秒前)の本処理において設定された和音候補M1, M2が今回の和音候補M1, M2として設定される(ステップS33)。ステップS32において和音候補数=1と判別された場合には今回のステップS30の実行では第1和音候補M1だけが設定されたので、第2和音候補M2は第1和音候補M1と同一の和音に設定される(ステップS34)。ステップS29~S34が和音候補検出手段に相当する。

【0029】

ステップS32において和音候補数>1と判別された場合には今回のステップS30の実行では第1及び第2和音候補M1, M2の両方が設定されたので、時刻、第1及び第2和音候補M1, M2が一時記憶メモリ6に記憶される(ステップS35)。一時記憶メモリ6には図8に示すように時刻、第1和音候補M1、第2和音候補M2が1組となって記憶される。時刻は0.2秒毎に増加するTで表される本処理実行回数である。そのTの順に第1及び第2和音候補M1, M2が記憶される。

【0030】

具体的には、一時記憶メモリ6に各和音候補を図8に示したように1バイトで記憶させるために、基本音(根音)とその属性との組み合わせが用いられる。基本音には平均律の12音が用いられ、属性にはメジャー{4, 3}、マイナー{3, 4}、セブンス候補{4, 6}及びディミニッシュセブンス(dim7)候補{3, 3}の和音の種類が用いられる。{}内は半音を1とした場合の3音の差である。本来、セブンス候補は{4, 3, 3}及びディミニッシュセブンス(dim7)候補{3, 3, 3}であるが、3音で示すために上記のように表示している。

【0031】

基本音の12音は図9(a)に示すように16ビット(16進表記)で表され、属性の和音の種類は同様に図9(b)に示すように16ビット(16進表記)で表される。その基本音の下位4ビットと属性の下位4ビットがその順に連結されて図9(c)に示すように8ビット(1バイト)として和音候補として用いられる。

ステップS35はステップS33又はS34を実行した場合にもその直後に実

行される。

【 0 0 3 2 】

ステップ S 3 5 の実行後、楽曲が終了したか否かが判別される（ステップ S 3 6）。例えば、デジタルオーディオ信号の入力がなくなった場合、或いは操作入力装置 2 からの楽曲の終了を示す操作入力があった場合には楽曲が終了したと判断される。これによって本処理が終了する。

楽曲の終了が判断されるまでは変数 T に 1 が加算され（ステップ S 3 7）、ステップ S 2 1 が再度実行される。ステップ S 2 1 は上記したように 0.2 秒間隔で実行され、前回の実行時から 0.2 秒が経過して再度実行される。

【 0 0 3 3 】

後処理においては、図 1 0 に示すように、一時記憶メモリ 6 から全ての第 1 及び第 2 和音候補が $M 1 (0) \sim M 1 (R)$ 及び $M 2 (0) \sim M 2 (R)$ として読み出される（ステップ S 4 1）。0 は開始時刻であり、開始時刻の第 1 及び第 2 和音候補が $M 1 (0)$ 及び $M 2 (0)$ である。R は最終時刻であり、最終時刻の第 1 及び第 2 和音候補が $M 1 (R)$ 及び $M 2 (R)$ である。読み出された第 1 和音候補 $M 1 (0) \sim M 1 (R)$ 及び第 2 和音候補 $M 2 (0) \sim M 2 (R)$ について平滑化が行われる（ステップ S 4 2）。この平滑化は和音の変化時点とは関係なく 0.2 秒間隔で和音候補を検出したことにより和音候補に含まれるノイズによる誤差を除去するために行われる。平滑化の具体的方法としては、3 つの連続する第 1 和音候補 $M 1 (t-1)$, $M 1 (t)$, $M 1 (t+1)$ について $M 1 (t-1) \neq M 1 (t)$ かつ $M 1 (t) \neq M 1 (t+1)$ の関係が成立するか否かが判別され、その関係が成立する場合には、 $M 1 (t+1)$ に $M 1 (t)$ は等しくされる。この判別は第 1 和音候補毎に行われる。第 2 和音候補についても同様の方法により平滑化は行われる。なお、 $M 1 (t+1)$ に $M 1 (t)$ を等しくするのではなく、逆に、 $M 1 (t+1)$ を $M 1 (t)$ に等しくしても良い。

【 0 0 3 4 】

平滑化後、第 1 及び第 2 和音候補の入れ替え処理が行われる（ステップ S 4 3）。一般的に 0.6 秒のような短い期間には和音が変化する可能性は低い。しかしながら、信号入力段の周波数特性及び信号入力時のノイズによって帯域データ $F'(T)$ 中の各音成分の周波数が変動することによって第 1 及び第 2 和音候補が

0. 6秒以内に入れ替わることが起きることがあり、これに対処するためにステップS43は行われる。第1及び第2和音候補が入れ替えの具体的方法としては、5つの連続する第1和音候補 $M1(t-2)$, $M1(t-1)$, $M1(t)$, $M1(t+1)$, $M1(t+2)$ 及びそれに対応する5つの連続する第2和音候補 $M2(t-2)$, $M2(t-1)$, $M2(t)$, $M2(t+1)$, $M2(t+2)$ についての次の如き判別が実行される。すなわち、 $M1(t-2)=M1(t+2)$, $M2(t-2)=M2(t+2)$, $M1(t-1)=M1(t)=M1(t+1)=M2(t-2)$ 及び $M2(t-1)=M2(t)=M2(t+1)=M1(t-2)$ の関係が成立するか否かが判別される。この関係が成立する場合には、 $M1(t-1)=M1(t)=M1(t+1)=M1(t-2)$ 及び $M2(t-1)=M2(t)=M2(t+1)=M2(t-2)$ が定められ、 $M1(t-2)$ と $M2(t-2)$ と間で和音の入れ替えが行われる。なお、 $M1(t-2)$ と $M2(t-2)$ との間で和音の入れ替えに代えて $M1(t+2)$ と $M2(t+2)$ との間で和音の入れ替えを行っても良い。また、 $M1(t-2)=M1(t+1)$, $M2(t-2)=M2(t+1)$, $M1(t-1)=M1(t)=M1(t+1)=M2(t-2)$ 及び $M2(t-1)=M2(t)=M2(t+1)=M1(t-2)$ の関係が成立するか否かが判別される。この関係が成立する場合には、 $M1(t-1)=M1(t)=M1(t-2)$ 及び $M2(t-1)=M2(t)=M2(t-2)$ が定められ、 $M1(t-2)$ と $M2(t-2)$ との間で和音の入れ替えが行われる。なお、 $M1(t-2)$ と $M2(t-2)$ との間で和音の入れ替えに代えて $M1(t+1)$ と $M2(t+1)$ との間で和音の入れ替えを行っても良い。

【0035】

ステップS41において読み出された第1和音候補 $M1(0) \sim M1(R)$ 及び第2和音候補 $M2(0) \sim M2(R)$ の各和音が、例えば、図11に示すように時間経過と共に変化する場合には、ステップS42の平均化を行うことによって図12に示すように修正される。更に、ステップS43の和音の入れ替えを行うことによって第1及び第2和音候補の和音の変化は図13に示すように修正される。なお、図11～図13は和音の時間変化を折れ線グラフとして示しており、縦軸は和音の種類に対応した位置となっている。

【0036】

ステップS43の和音の入れ替え後の第1和音候補 $M1(0) \sim M1(R)$ のうちの

和音が変化した時点 t の $M1(t)$ 及び第 2 和音候補 $M2(0) \sim M2(R)$ のうちの和音が変化した時点 t の $M2(t)$ が各々検出され (ステップ S 4 4)、その検出された時点 t (4 バイト) 及び和音 (4 バイト) が第 1 及び第 2 和音候補毎にデータ蓄積装置 5 に記憶される (ステップ S 4 5)。ステップ S 4 5 で記憶される 1 楽曲分のデータが和音進行楽曲データである。かかるステップ S 4 1 \sim S 4 5 が平滑化手段に相当する。

【 0 0 3 7 】

ステップ S 4 3 の和音の入れ替え後の第 1 和音候補 $M1(0) \sim M1(R)$ 及び第 2 和音候補 $M2(0) \sim M2(R)$ の和音が図 1 4 (a) に示すように時間経過と共に変化する場合には、変化時点の時刻と和音がデータとして抽出される。図 1 4 (b) が第 1 和音候補の変化時点のデータ内容であり、F, G, D, B \flat , F が和音であり、それらは 16 進データとして 0x08, 0x0A, 0x05, 0x01, 0x08 と表される。変化時点 t の時刻は $T1(0)$, $T1(1)$, $T1(2)$, $T1(3)$, $T1(4)$ である。また、図 1 4 (c) が第 2 和音候補の変化時点のデータ内容であり、C, B \flat , F \sharp m, B \flat , C が和音であり、それらは 16 進データとして 0x03, 0x01, 0x29, 0x01, 0x03 と表される。変化時点 t の時刻は $T2(0)$, $T2(1)$, $T2(2)$, $T2(3)$, $T2(4)$ である。図 1 4 (b) 及び図 1 4 (c) に示したデータ内容は楽曲の識別情報と共にデータ蓄積装置 5 には、ステップ S 4 5 においては図 1 4 (d) に示すような形式で 1 ファイルとして記憶される。

【 0 0 3 8 】

異なる楽曲音を示すオーディオ信号について上記した和音分析動作を繰り返すことによりデータ蓄積装置 5 には複数の楽曲毎のファイルとして和音進行楽曲データが蓄積されることになる。なお、データ蓄積装置 4 にはデータ蓄積装置 5 の和音進行楽曲データに対応した PCM 信号からなる楽曲データが蓄積される。

ステップ S 4 4 において第 1 和音候補のうちの和音が増えた時点の第 1 和音候補及び第 2 和音候補のうちの和音が増えた時点の第 2 和音候補が各々検出され、それが最終的な和音進行楽曲データとなるので、MP 3 のような圧縮データに比べても 1 楽曲当たりの容量を小さくすることができ、また、各楽曲のデータ

を高速処理することができる。

【 0 0 3 9 】

また、データ蓄積装置 5 に書き込まれた和音進行楽曲データは、実際の楽曲と時間的に同期した和音データとなるので、第 1 和音候補のみ、或いは第 1 和音候補と第 2 和音候補との論理和出力を用いて実際に和音を楽曲再生装置 1 0 によって生成すれば、楽曲の伴奏が可能となる。

次に、データ蓄積装置 5 に和音進行楽曲データとして蓄積された楽曲の構造を検出する楽曲構造検出動作について説明する。楽曲構造検出動作は和音進行比較装置 7 及び繰り返し構造検出装置 8 によって実行される。

【 0 0 4 0 】

楽曲構造検出動作においては、図 1 5 に示すように、楽曲構造検出対象の楽曲の第 1 和音候補 $M1(0) \sim M1(a-1)$ 及び第 2 和音候補 $M2(0) \sim M2(b-1)$ が蓄積手段であるデータ蓄積装置 5 から読み出される (ステップ S 5 1)。その楽曲構造検出対象の楽曲は例えば、操作入力装置 2 の操作によって指定される。a は第 1 和音候補の総数であり、b は第 2 和音候補の総数である。また、仮想データとして各々 K 個の第 1 和音候補 $M1(a) \sim M1(a+K-1)$ 及び第 2 和音候補 $M2(b) \sim M2(b+K-1)$ が用意される (ステップ S 5 2)。ここで、 $a < b$ のとき仮想データの第 1 及び第 2 和音候補各々の和音総数 P は a に等しく、 $a \geq b$ のとき和音総数 P は b に等しい。仮想データは第 1 和音候補 $M1(0) \sim M1(a-1)$ 及び第 2 和音候補 $M2(0) \sim M2(b-1)$ の後に付加される。

【 0 0 4 1 】

読み出された第 1 和音候補 $M1(0) \sim M1(P-1)$ に対して第 1 和音差分値 $MR1(0) \sim MR1(P-2)$ が計算される (ステップ S 5 3)。第 1 和音差分値は、 $MR1(0) = M1(1) - M1(0)$, $MR1(1) = M1(2) - M1(1)$, …… , $MR1(P-2) = M1(P-1) - M1(P-2)$ の如く計算される。この計算では第 1 和音差分値 $MR1(0) \sim MR1(P-2)$ 各々が 0 より小であるか否かを判別し、0 より小の第 1 和音差分値には 1 2 を加算することが行われる。また、第 1 和音差分値 $MR1(0) \sim MR1(P-2)$ 各々には和音変化後の和音属性 $MA1(0) \sim MA1(P-2)$ が付加される。読み出された第 2 和音候補 $M2(0) \sim M2(P-1)$ に対しても第 2 和音差分値 $MR2$

(0)~MR 2 (P-2)が計算される(ステップS 5 4)。第2和音差分値は、MR 2 (0)=M 2 (1)-M 2 (0), MR 2 (1)=M 2 (2)-M 2 (1), …… , MR 2 (P-2)=M 2 (P-1)-M 2 (P-2)の如く計算される。この計算においても第2和音差分値MR 2 (0)~MR 2 (P-2)各々が0より小であるか否かを判別し、0より小の第2和音差分値には1 2を加算することが行われる。また、第2和音差分値MR 2 (0)~MR 2 (P-2)各々には和音変化後の和音属性MA 2 (0)~MA 2 (P-2)が付加される。なお、和音属性MA 1 (0)~MA 1 (P-2), MA 2 (0)~MA 2 (P-2)には図9 (b)に示した数値が用いられる。

【0 0 4 2】

図1 6はステップS 5 3及びS 5 4の動作例を説明している。すなわち、和音候補がAm 7, Dm, C, F, Em, F, B b #の列である場合に、和音差分値は5, 1 0, 5, 1 1, 1, 5となり、和音変化後の和音属性は0 x 0 2, 0 x 0 0, 0 x 0 0, 0 x 0 2, 0 x 0 0, 0 x 0 0となる。なお、和音変化後の和音属性がセブンスの場合にはそれに代えてメジャーとしている。セブンスを用いてもその比較演算結果への影響が小さいので、演算量を削減するためである。

【0 0 4 3】

ステップS 5 4の実行後、カウンタ値cが0に初期化される(ステップS 5 5)。そして、第1和音候補M 1 (0)~M 1 (P-1)及び第2和音候補M 2 (0)~M 2 (P-1)各々のうちのc番目からK個(例えば、2 0)の和音候補(部分楽曲データ)が抽出される(ステップS 5 6)。すなわち、第1和音候補M 1 (c)~M 1 (c+K-1)及び第2和音候補M 2 (c)~M 2 (c+K-1)が抽出される。M 1 (c)~M 1 (c+K-1)=U 1 (0)~U 1 (K-1)とし、M 2 (c)~M 2 (c+K-1)=U 2 (0)~U 2 (K-1)とする。図1 7は処理対象の和音進行データのM 1 (0)~M 1 (P-1), M 2 (0)~M 2 (P-1)及び仮想データに対するU 1 (0)~U 1 (K-1), U 2 (0)~U 2 (K-1)の関係を示している。

【0 0 4 4】

ステップ S 5 6 の実行後、部分楽曲データの第 1 和音候補 $U 1 (0) \sim U 1 (K-1)$ に対して第 1 和音差分値 $U R 1 (0) \sim U R 1 (K-2)$ が計算される (ステップ S 5 7)。ステップ S 5 7 の第 1 和音差分値は、 $U R 1 (0) = U 1 (1) - U 1 (0)$, $U R 1 (1) = U 1 (2) - U 1 (1)$, \dots , $U R 1 (K-2) = U 1 (K-1) - U 1 (K-2)$ の如く計算される。この計算では第 1 和音差分値 $U R 1 (0) \sim U R 1 (K-2)$ 各々が 0 より小であるか否かを判別し、0 より小の第 1 和音差分値には 1 2 を加算することが行われる。また、第 1 和音差分値 $U R 1 (0) \sim U R 1 (K-2)$ 各々には和音変化後の和音属性 $U A 1 (0) \sim U A 1 (K-2)$ が付加される。また、部分楽曲データの第 2 和音候補 $U 2 (0) \sim U 2 (K-1)$ に対しても第 2 和音差分値 $U R 2 (0) \sim U R 2 (K-2)$ が計算される (ステップ S 5 8)。第 2 和音差分値は、 $U R 2 (0) = U 2 (1) - U 2 (0)$, $U R 2 (1) = U 2 (2) - U 2 (1)$, \dots , $U R 2 (K-2) = U 2 (K-1) - U 2 (K-2)$ の如く計算される。この計算においても第 2 和音差分値 $U R 2 (0) \sim U R 2 (K-2)$ 各々が 0 より小であるか否かを判別し、0 より小の第 2 和音差分値には 1 2 を加算することが行われる。また、第 2 和音差分値 $U R 2 (0) \sim U R 2 (K-2)$ 各々には和音変化後の和音属性 $U A 2 (0) \sim U A 2 (K-2)$ が付加される。

【 0 0 4 5 】

ステップ S 5 3 にて得られた第 1 和音差分値 $M R 1 (0) \sim M R 1 (K-2)$ 及び和音属性 $M A 1 (0) \sim M A 1 (K-2)$ と、ステップ S 5 7 にて得られた c 番目から K 個の第 1 和音候補 $U R 1 (0) \sim U R 1 (K-2)$ 及び和音属性 $U A 1 (0) \sim U A 1 (K-2)$ と、ステップ S 5 8 にて得られた c 番目から K 個の第 2 和音候補 $U R 2 (0) \sim U R 2 (K-2)$ 及び和音属性 $U A 2 (0) \sim U A 2 (K-2)$ とに応じて相互相関演算が行われる (ステップ S 5 9)。相互相関演算では相関係数 $C O R (t)$ が次式 (3) の如く算出される。相関係数 $C O R (t)$ が小さいほど類似性が高いことを示す。

$$\begin{aligned}
 C O R (t) = & \sum 10 (|M R 1 (t+k) - U R 1 (k')| + |M A 1 (t+k) - U A 1 (k')| \\
 & + |W M 1 (t+k+1) / W M 1 (t+k) - W U 1 (k'+1) / W U 1 (k')|) \\
 & + \sum 10 (|M R 1 (t+k) - U R 2 (k')| + |M A 1 (t+k) - U A 2 (k')| \\
 & + |W M 1 (t+k+1) / W M 1 (t+k) - W U 2 (k'+1) / W U 2 (k')|) \quad \dots (3)
 \end{aligned}$$

ただし、 $W U 1 ()$, $W M 1 ()$, $W U 2 ()$ は各和音が維持される時間幅、 $t = 0 \sim P - 1$ 、 \sum 演算は $k = 0 \sim K - 2$ 及び $k' = 0 \sim K - 2$ である。

【 0 0 4 6 】

ステップ S 5 9 の相関係数 $COR(t)$ は t が $0 \sim P-1$ の範囲で各々算出される。また、ステップ S 5 9 の相関係数 $COR(t)$ の演算では飛び越し処理が行われる。飛び越し処理においては、 $(MR1(t+k+k1) - UR1(k'+k2))$ 又は $(MR1(t+k+k1) - UR2(k'+k2))$ の最小値が検出される。 $k1$ 及び $k2$ 各々は $0 \sim 2$ までのいずれかの整数である。すなわち、 $k1$ 及び $k2$ 各々を $0 \sim 2$ までの範囲で変化させて $(MR1(t+k+k1) - UR1(k'+k2))$ 又は $(MR1(t+k+k1) - UR2(k'+k2))$ の最小値となるときの検出される。そのときの $k+k1$ が新たな k に、 $k'+k2$ が新たな k' とされる。その後、式(3)に応じて相関係数 $COR(t)$ が算出される。

【 0 0 4 7 】

更に、各時点の和音から変化後の和音が処理対象の和音進行楽曲データ及びその和音進行楽曲データの c 番目から K 個の部分楽曲データが C 及び A_m のいずれであっても、或いは C_m 及び E_b のいずれであっても同一とみなす。すなわち、変化後の和音が関係調の和音であれば、上記の式の $|MR1(t+k) - UR1(k')| + |MA1(t+k) - UA1(k')| = 0$ 又は $|MR1(t+k) - UR2(k')| + |MA1(t+k) - UA2(k')| = 0$ である。例えば、和音 F から一方のデータが 7 度差でメジャーに変化し、他方のデータが 4 度差でマイナーに変化した場合には同一とし、また和音 F から一方のデータが 7 度差でマイナーに変化し、他方のデータが 10 度差でメジャーに変化した場合にも同一として処理される。

【 0 0 4 8 】

更に、ステップ S 5 4 にて得られた第 2 和音差分値 $MR2(0) \sim MR2(K-2)$ 及び和音属性 $MA2(0) \sim MA2(K-2)$ と、ステップ S 5 7 にて得られた c 番目から K 個の第 1 和音候補 $UR1(0) \sim UR1(K-2)$ 及び和音属性 $UA1(0) \sim UA1(K-2)$ と、ステップ S 5 8 にて得られた c 番目から K 個の第 2 和音候補 $UR2(0) \sim UR2(K-2)$ 及び和音属性 $UA2(0) \sim UA2(K-2)$ とに応じて相互相関演算が行われる(ステップ S 6 0)。相互相関演算では相関係数 $COR'(t)$ が次式(4)の如く算出される。相関係数 $COR'(t)$ が小さいほど類似性が高いことを示す。

$$COR'(t) = \sum 10(|MR2(t+k) - UR1(k')| + |MA2(t+k) - UA1(k')|$$

$$\begin{aligned}
& + |WM 2 (t+k+1)/WM 2 (t+k) - WU 1 (k'+1)/WU 1 (k')|) \\
& + \Sigma 10(|MR 2 (t+k) - UR 2 (k')| + |MA 2 (t+k) - UA 2 (k')| \\
& + |WM 2 (t+k+1)/WM 2 (t+k) - WU 2 (k'+1)/WU 2 (k')|) \dots\dots (4)
\end{aligned}$$

ただし、 $WU 1 ()$ 、 $WM 2 ()$ 、 $WU 2 ()$ は各和音が維持される時間幅、 $t = 0 \sim P - 1$ 、 Σ 演算は $k = 0 \sim K - 2$ 及び $k' = 0 \sim K - 2$ である。

【 0 0 4 9 】

ステップ S 6 0 の相関係数 $COR'(t)$ は t が $0 \sim P - 1$ の範囲で各々算出される。また、ステップ S 6 0 の相関係数 $COR(t)$ の演算では上記のステップ S 5 9 と同様に飛び越し処理が行われる。飛び越し処理においては、 $(MR 2 (t+k+k_1) - UR 1 (k'+k_2))$ 又は $(MR 2 (t+k+k_1) - UR 2 (k'+k_2))$ の最小値が検出される。 k_1 及び k_2 各々は $0 \sim 2$ までのいずれかの整数である。すなわち、 k_1 及び k_2 各々を $0 \sim 2$ までの範囲で変化させて $(MR 2 (t+k+k_1) - UR 1 (k'+k_2))$ 又は $(MR 2 (t+k+k_1) - UR 2 (k'+k_2))$ の最小値となるときの $k+k_1$ が新たな k に、 $k'+k_2$ が新たな k' とされる。その後、式(4)に応じて相関係数 $COR'(t)$ が算出される。

【 0 0 5 0 】

更に、各時点の和音から変化後の和音が処理対象の和音進行楽曲データ及び部分楽曲データが C 及び A m のいずれであっても、或いは C m 及び E b のいずれであっても同一とみなす。すなわち、変化後の和音が関係調の和音であれば、上記の式の $|MR 2 (t+k) - UR 1 (k')| + |MA 2 (t+k) - UA 1 (k')| = 0$ 又は $|MR 2 (t+k) - UR 2 (k')| + |MA 2 (t+k) - UA 2 (k')| = 0$ である。

【 0 0 5 1 】

図 1 8 (a) は処理対象の和音進行楽曲データとその部分楽曲データとの関係を示している。部分楽曲データは t の進行に従って処理対象の和音進行楽曲データとの比較部分が変化する。図 1 8 (b) は相関係数 $COR(t)$ 又は $COR'(t)$ の変化を示している。ピーク波形部分が類似性が高い部分である。

図 1 8 (c) は処理対象の和音進行楽曲データとその部分楽曲データとの相互相関演算における、各和音が維持される時間幅 $WU(1) \sim WU(5)$ 、飛び越し処理部分及び関係調の部分を示している。処理対象の和音進行楽曲データと部分楽曲デ

ータとの間の矢印線は同一和音を示している。その矢印線のうちの同一時間にな
い傾いた矢印線で結ばれた和音は、飛び越し処理で検出された和音である。また
、矢印線が波線になっているものは関係調の和音である。

【 0 0 5 2 】

ステップ S 5 9 及び S 6 0 で算出された相関係数 $COR(t)$ 及び $COR'(t)$ は
加算されて合計相関係数 $COR(c,t)$ が算出される (ステップ S 6 1)。すなわ
ち、 $COR(c,t)$ は次式 (5) に示すように算出される。

$$COR(c,t) = COR(t) + COR'(t) \quad t = 0 \sim P-1 \quad \dots\dots (5)$$

図 1 9 (a)～(f) は処理対象の和音進行楽曲データが示す楽曲中のフレーズ (和
音進行列) と、部分楽曲データが示すフレーズと、合計相関係数 $COR(c,t)$ と
の関係を示している。和音進行楽曲データが示す楽曲中のフレーズは図示しない
イントロ I の後の曲の流れ順に A, B, C, A', C', D, C" であり、A と A'
とが同一フレーズ、また C と C' と C" とが同一フレーズとする。図 1 9 (a) では
、部分楽曲データの先頭にフレーズ A が位置している場合にであり、 $COR(c,t)$
) は和音進行楽曲データのフレーズ A と A' とに対応した時点で□で示すピーク値
を生成する。図 1 9 (b) では、部分楽曲データの先頭にフレーズ B が位置してい
る場合にであり、 $COR(c,t)$ は和音進行楽曲データのフレーズ B だけに対応し
た時点で×で示すピーク値を生成する。図 1 9 (c) では、部分楽曲データの先頭
にフレーズ C が位置している場合にであり、 $COR(c,t)$ は和音進行楽曲データ
のフレーズ C, C', C" の各々に対応した時点で○で示すピーク値を生成する
。図 1 9 (d) では、部分楽曲データの先頭にフレーズ A' が位置している場合にで
あり、 $COR(c,t)$ は和音進行楽曲データのフレーズ A, A' の各々に対応した時
点で□で示すピーク値を生成する。図 1 9 (e) では、部分楽曲データの先頭にフ
レーズ C' が位置している場合にであり、 $COR(c,t)$ は和音進行楽曲データのフ
レーズ C, C', C" の各々に対応した時点で○で示すピーク値を生成する。図
1 9 (f) では、部分楽曲データの先頭にフレーズ C" が位置している場合にであ
り、 $COR(c,t)$ は和音進行楽曲データのフレーズ C, C', C" の各々に対応し
た時点で○で示すピーク値を生成する。

【 0 0 5 3 】

ステップ S 6 1 の実行後、カウンタ値 c に 1 が加算され（ステップ S 6 2）、そのカウンタ値 c が $P - 1$ より大であるか否かが判別される（ステップ S 6 3）。 $c \leq P - 1$ であるならば、処理対象の和音進行楽曲データ全てに亘って相関係数 $COR(c, t)$ が算出されていない。よって、ステップ S 5 6 に戻って上記のステップ S 5 6 ~ S 6 3 の動作が繰り返される。

【 0 0 5 4 】

$c > P - 1$ であるならば、 $COR(c, t)$ 、すなわち $COR(0, 0) \sim COR(P - 1, P - 1)$ のピーク値が検出され、そのピーク値の検出時の c , t について $COR_PEAK(c, t) = 1$ が設定され、ピーク値でないときの c , t について $COR_PEAK(c, t) = 0$ が設定される（ステップ S 6 4）。 $COR(c, t)$ が所定値を越えた部分の最高値をピーク値とする。ステップ 6 4 によって $COR_PEAK(c, t)$ の列が形成される。次に、この $COR_PEAK(c, t)$ 列において、 t が $0 \sim P - 1$ の各々の $COR_PEAK(c, t)$ の合計値がピーク数 $PK(t)$ として算出される（ステップ S 6 5）。 $PK(0) = COR_PEAK(0, 0) + COR_PEAK(1, 0) + \dots + COR_PEAK(P - 1, 0)$, $PK(1) = COR_PEAK(0, 1) + COR_PEAK(1, 1) + \dots + COR_PEAK(P - 1, 1)$, \dots , $PK(P - 1) = COR_PEAK(0, P - 1) + COR_PEAK(1, P - 1) + \dots + COR_PEAK(P - 1, P - 1)$ である。ピーク数 $PK(0) \sim PK(P - 1)$ のうちの連続する 2 以上の同一数の範囲が同一フレーズ範囲として区分けされ、それに基づいて楽曲構造データがデータ蓄積装置 5 に保存される（ステップ S 6 6）。例えば、ピーク数 $PK(t)$ が 2 である場合には、楽曲中で 2 回繰り返しが行われるフレーズとなり、ピーク数 $PK(t)$ が 3 である場合には、楽曲中で 3 回繰り返しが行われるフレーズとなる。同一フレーズの範囲のピーク数 $PK(t)$ は同一値となる。ピーク数 $PK(t)$ が 1 である場合には、繰り返しが無いフレーズを示すことになる。

【 0 0 5 5 】

図 2 0 は図 1 9 (a) ~ (f) に示したフレーズ I, A, B, C, A', C', D, C" を有する楽曲についてのピーク数 $PK(t)$ と、相関係数 $COR(c, t)$ の算出結果からピーク値が得られた位置 $COR_PEAK(c, t)$ とを示している。 $COR_PEAK(c, t)$ はマトリックスで表示しており、横軸が和音数 $t = 0 \sim P - 1$ で

あり、縦軸が部分楽曲データの開始位置である $c = 0 \sim P - 1$ を示している。ドット部分が $COR(c, t)$ がピーク値を得た $COR_PEAK(c, t) = 1$ に対応した位置である。対角線上は同一データ同士の自己相関をとったことになるので、ドット列となる。対角線以外の部分に現れるドット列が繰り返しの和音進行によるフレーズに対応する。図 19 (a)～(f) に対応して \times は 1 回だけのフレーズ I, B, D に対応し、 \bigcirc は 3 回の繰り返しフレーズ C, C', C" に対応し、 \square は 2 回の繰り返しフレーズ A, A' に対応する。ピーク数 $PK(t)$ はフレーズ I, A, B, C, A', C', D, C" に対応して 1, 2, 1, 3, 2, 3, 1, 3 となる。これが結果として楽曲構造を示すことになる。

【0056】

楽曲構造データは図 21 に示すようなフォーマットを有している。各フレーズの開始時刻情報及び終了時刻情報には図 14 (c) に示した和音進行楽曲データ $T(t)$ が用いられる。

また、楽曲構造検出結果が表示装置 9 に表示される (ステップ S 6 7)。楽曲構造検出結果の表示画面は図 22 に示すように楽曲中の各繰り返しフレーズ部分の選択ができるようにされている。この表示画面によって選択された繰り返しフレーズ部分又は繰り返し回数が最も多いフレーズ部分に対応する楽曲データがデータ蓄積装置 4 から読み出されて楽曲再生装置 10 に供給される (ステップ S 6 8)。これにより、楽曲再生装置 10 は供給された楽曲データを順次再生し、それがデジタル信号としてデジタル／アナログ変換装置 11 に供給される。デジタル／アナログ変換装置 11 においてアナログオーディオ信号に変換された後、スピーカ 15 から繰り返しフレーズ部分の再生音が出力されることになる。

【0057】

よって、利用者は処理対象の楽曲の構造を表示画面から知ることができると共に、その楽曲のうちの選択した繰り返しフレーズ部分又は繰り返し回数が最も多いフレーズ部分を容易に聴取することができる。

上記の楽曲構造検出動作のステップ S 5 6 が部分楽曲データ生成手段に対応し、ステップ S 5 7～S 6 3 が類似度 (相関係数 $COR(C, t)$) を算出する比較手段に相当し、ステップ S 6 4 が和音位置検出手段に相当し、ステップ S 6 5～

S 6 8 が出力手段に相当する。

【 0 0 5 8 】

上記した飛び越し処理及び関係調処理は、和音の変化前後の差分値の演算の際に処理対象の和音進行楽曲データがアナログ信号に基づいて作成された場合における外部雑音や入力装置の周波数特性の影響を排除するため、或いは1番と2番とでは同一フレーズであってもリズムや旋律の変化があったり、又は転調が行われている場合にはデータ間の和音の位置や属性が完全に一致しないことが起きるので、それを防止するために行われる。すなわち、一時的に和音進行が異なっても一定時間幅内で和音進行の傾向が類似していることを検出することができるので、リズムや旋律の変化があったり、又は転調が行われている場合でもその影響を受けることなく、同一フレーズであるか否かを正確に判別することができる。更に、飛び越し処理及び関係調処理を施すことによってその施した部分以外の相互関連演算においても正確な類似度を求めることができる。

【 0 0 5 9 】

また、上記した実施例においては、PCMデータ形式の楽曲データに対して作用することを前提としているが、ステップS 2 8の処理において楽曲に含まれる音符列が分かっているれば、楽曲データとしてMIDIデータを用いることもできる。更に、上記した実施例のシステムを応用すれば、楽曲を構成する繰り返し回数の多いフレーズ部分だけを順に再生する、例えば、ハイライト再生システムを実現することも容易に可能である。

【 0 0 6 0 】

図2 3は本発明の他の実施例を示している。図2 3の楽曲処理システムにおいては、図1のシステム中の和音解析装置3、一時記憶メモリ6、和音進行比較装置7及び繰り返し構造検出装置8がコンピュータ2 1によって形成されている。コンピュータ2 1は記憶装置2 2に記憶されたプログラムに応じて上記の和音解析動作及び楽曲構造検出動作を実行する。記憶装置2 2はハードディスクドライブに限らず、記録媒体のドライブ装置でも良い。その記録媒体のドライブ装置の場合には記録媒体に和音進行楽曲データを書き込むようにしても良い。

【 0 0 6 1 】

以上のように、本発明によれば、和音進行楽曲データ中の各和音の位置から連続する所定数の和音からなる部分楽曲データを生成する部分楽曲データ生成手段と、部分楽曲データ各々と和音進行楽曲データとを和音進行楽曲データ中の各和音の位置から和音変化時の和音の根音変化量と変化後の和音の属性とについて比較して複数の楽曲毎の類似度を算出する比較手段と、部分楽曲データ毎に比較手段によって算出された類似度各々に応じて類似度が所定値より高いピーク値となった和音進行楽曲データ中の和音の位置を検出する和音位置検出手段と、和音進行楽曲データ中の和音の位置毎に部分楽曲データ全てについて類似度が所定値より高いピーク値となった回数を算出し、その和音の位置毎の算出回数に応じて楽曲構造を示す検出出力を生成する出力手段と、を備えたことにより、繰り返し部分を含む楽曲の構造を簡単な構成で適切に検出することができる。

【図面の簡単な説明】

【図 1】

本発明を適用した楽曲処理システムの構成を示すブロック図である。

【図 2】

周波数誤差検出動作を示すフローチャートである。

【図 3】

A 音を 1.0 とした場合の 1 2 音及び 1 オクターブ高い A 音各々の周波数比を示す図である。

【図 4】

和音解析動作の本処理を示すフローチャートである。

【図 5】

帯域データの各音成分の強度レベル例を示す図である。

【図 6】

帯域データの各音成分の強度レベル例を示す図である。

【図 7】

4 音からなる和音に対する 3 音からなる和音への変換を示す図である。

【図 8】

一時記憶メモリへの記録フォーマットを示す図である。

【図 9】

基本音及び和音の属性の表記方法、並びに和音候補の表記方法を示す図である。

【図 1 0】

和音解析動作の後処理を示すフローチャートである。

【図 1 1】

平滑化処理前の第 1 及び第 2 和音候補の時間変化を示す図である。

【図 1 2】

平滑化処理後の第 1 及び第 2 和音候補の時間変化を示す図である。

【図 1 3】

入れ替え処理後の第 1 及び第 2 和音候補の時間変化を示す図である。

【図 1 4】

和音進行楽曲データの作成方法及びそのフォーマットを示す図である。

【図 1 5】

楽曲構造検出動作を示すフローチャートである。

【図 1 6】

和音変化の和音差分値及び変化後の属性の例を示す図である。

【図 1 7】

仮想データを含む和音進行楽曲データと部分楽曲データとの関係を示す図である。

【図 1 8】

相互相関演算時の和音進行楽曲データと部分楽曲データとの関係、相関係数 $COR(c, t)$ の変化、並びに各和音が維持される時間幅、飛び越し処理部分及び関係調の部分を示している。

【図 1 9】

部分楽曲データに含まれるフレーズと和音進行楽曲データに含まれるフレーズ列とに応じた相関係数 $COR(c, t)$ の変化を示す図である。

【図 2 0】

図 1 9 に示したフレーズ列を有する楽曲についてのピーク数 $PK(t)$ と、ピー

ク値が得られた位置COR__PEAK(c,t)とを示す図である。

【図 2 1】

楽曲構造データのフォーマットを示す図である。

【図 2 2】

表示装置の表示例を示す図である。

【図 2 3】

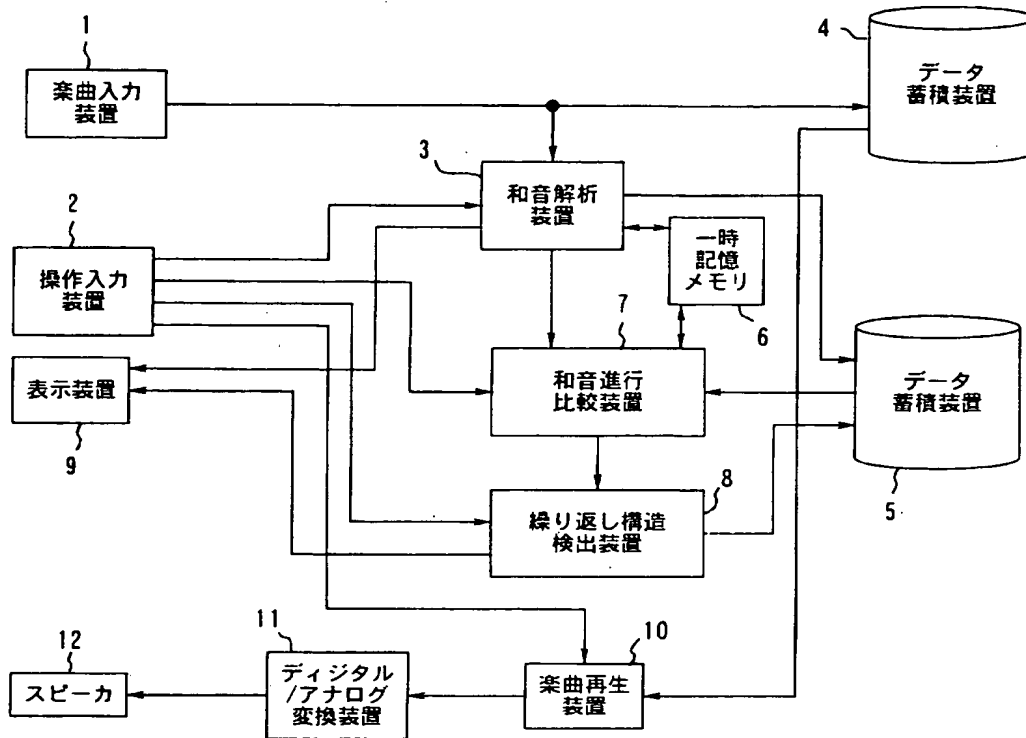
本発明の他の実施例として楽曲処理システムの構成を示すブロック図である。

【符号の説明】

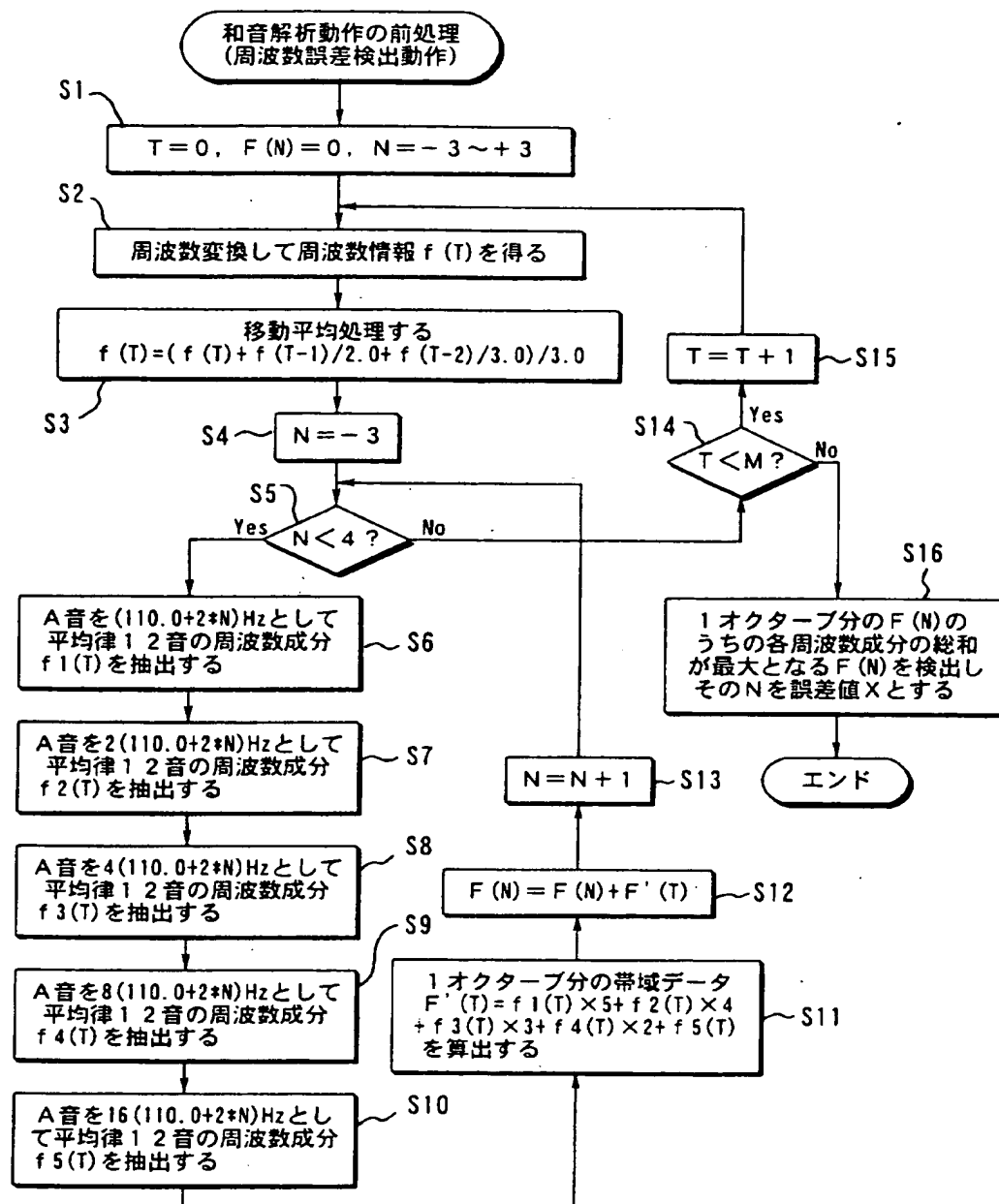
- 3 和音解析装置
- 4, 5 データ蓄積装置
- 7 和音進行比較装置
- 8 繰り返し構造検出装置
- 10 楽曲再生装置
- 21 コンピュータ

【書類名】 図面

【図 1】



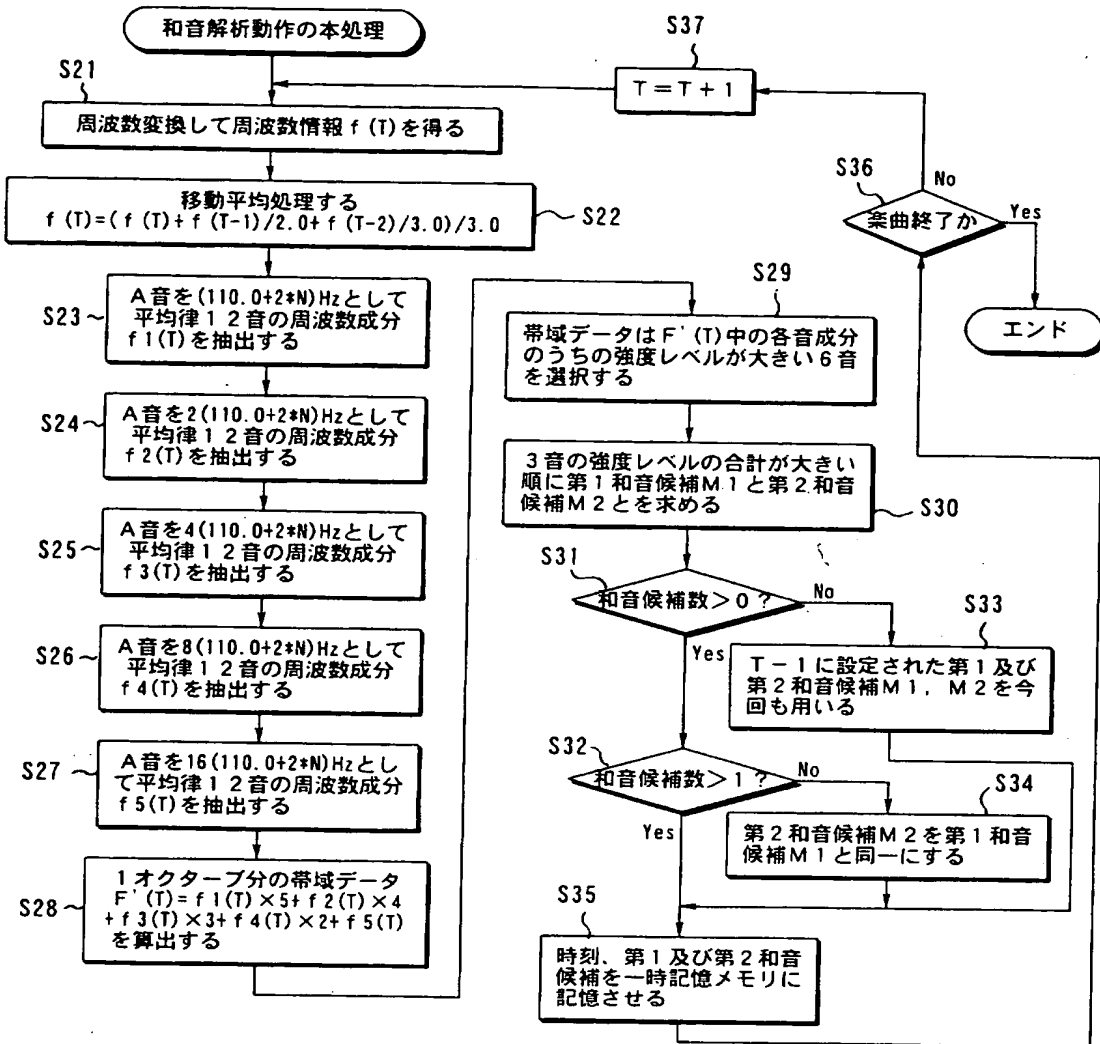
【図 2】



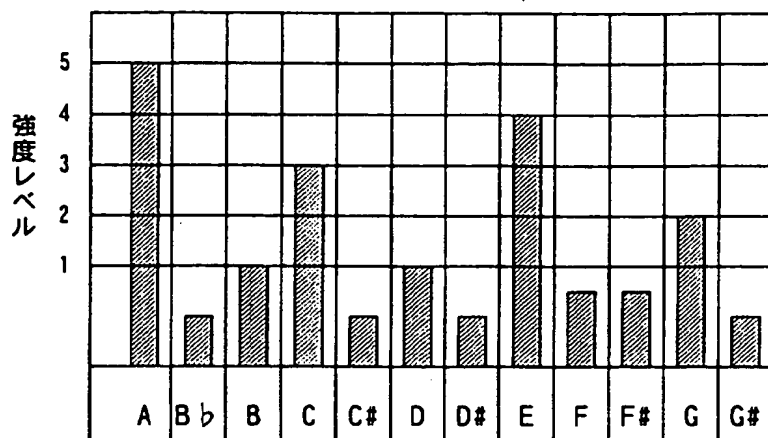
【図 3】

A	1.0
A #	AX1.059163
B	AX1.124622
C	AX1.189207
C #	AX1.259921
D	AX1.334840
D #	AX1.414214
E	AX1.498307
F	AX1.587401
F #	AX1.681793
G	AX1.781797
G #	AX1.887749
A	AX2.0

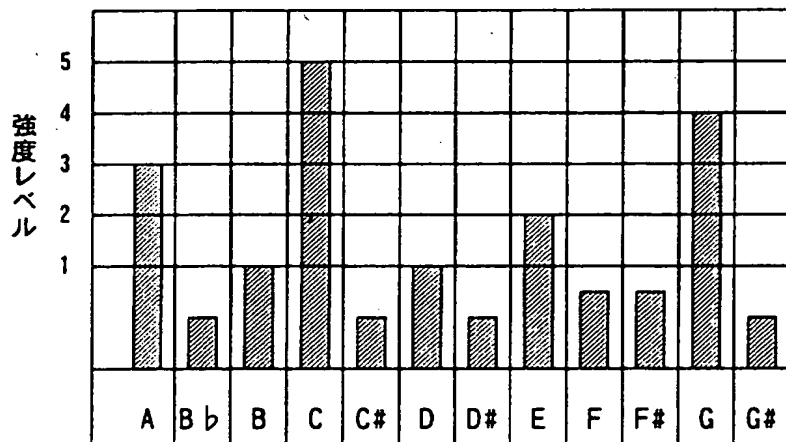
【図 4】



【図 5】



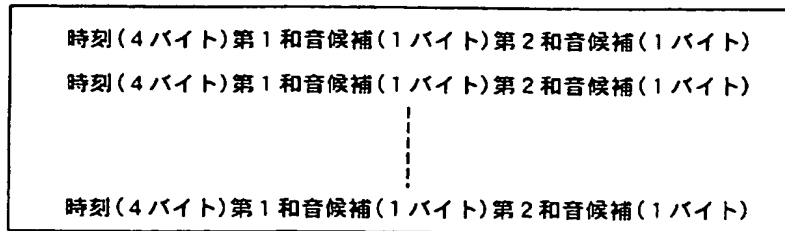
【図 6】



【図 7】

$A+C+E+G(Am7)$
 $\rightarrow A+C+E(Am) \mid C+E+G(C)$
 $A+C+E^b + F\#(Adim7)$
 $\rightarrow A+C+E^b (Adim7候補) \mid C+E^b + F\#(Cdim7候補)$
 $A+C\#+E+G(A7)$
 $\rightarrow A+C\#+E(A) \mid A+C\#+G(A7候補) \mid C\#+E+G(C\#dim7候補)$
 $A+C+E+G+B(Am7.9)$
 $\rightarrow A+C+E(Am) \mid C+E+G(C) \mid E+G+B(Em)$

【図 8】



【図 9】

(a)

基本音	
A	0 x 0 0
A #	0 x 0 1
B	0 x 0 2
C	0 x 0 3
C #	0 x 0 4
D	0 x 0 5
D #	0 x 0 6
E	0 x 0 7
F	0 x 0 8
F #	0 x 0 9
G	0 x 0 A
G #	0 x 0 B

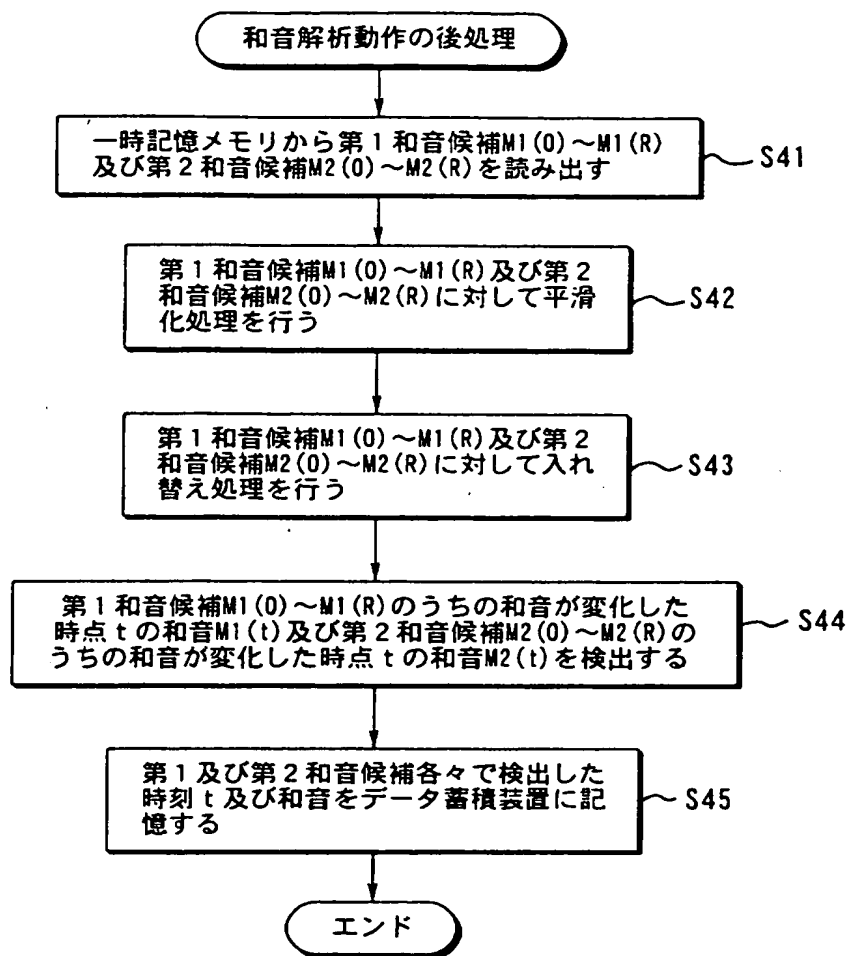
(b)

属性	
メジャー	0 x 0 0
マイナー	0 x 0 2
セブンス候補	0 x 0 1
Dim7 候補	0 x 0 3

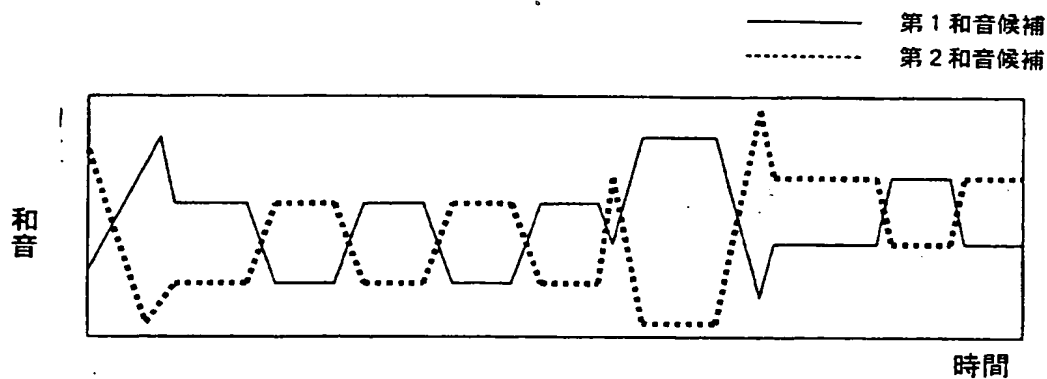
(c)

属性(下位 4 ビット) + 基本音(下位 4 ビット)

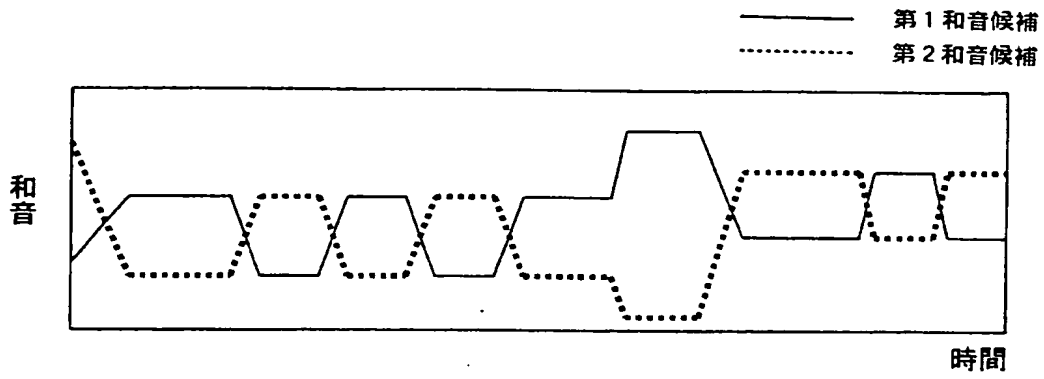
【図 1 0】



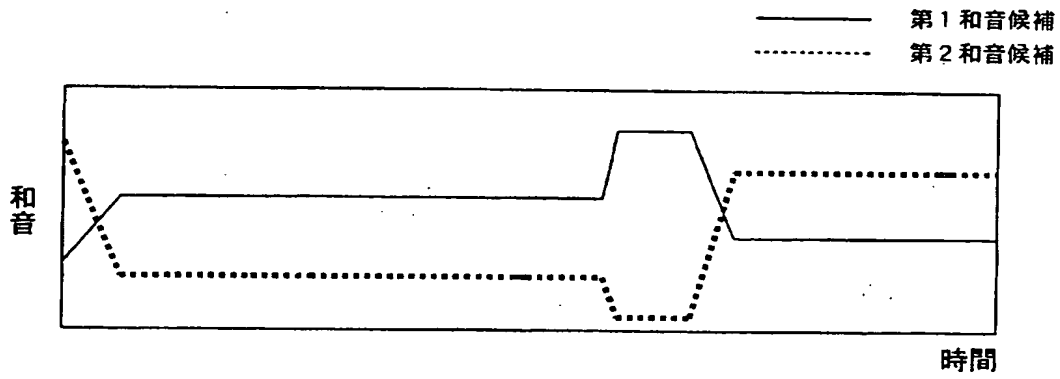
【図 1 1】



【図 1 2】

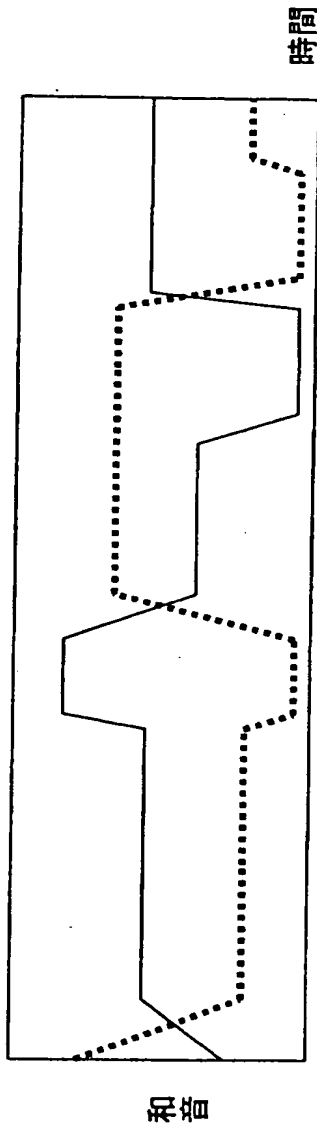


【図 1 3】



【図 1 4】

—— 第1和音候補
 第2和音候補



(a)

(b) 第1和音候補
 の変化推移

F	G	D	B \flat	F
0x08	0x0A	0x05	0x01	0x08
←時刻(T1(0))	T1(1)	T1(2)	T1(3)	T1(4)

C	B \flat	F \sharp m	B \flat	C
0x03	0x01	0x29	0x01	0x03
←時刻(T2(0))	T2(1)	T2(2)	T2(3)	T2(4)

(c) 第2和音候補
 の変化推移

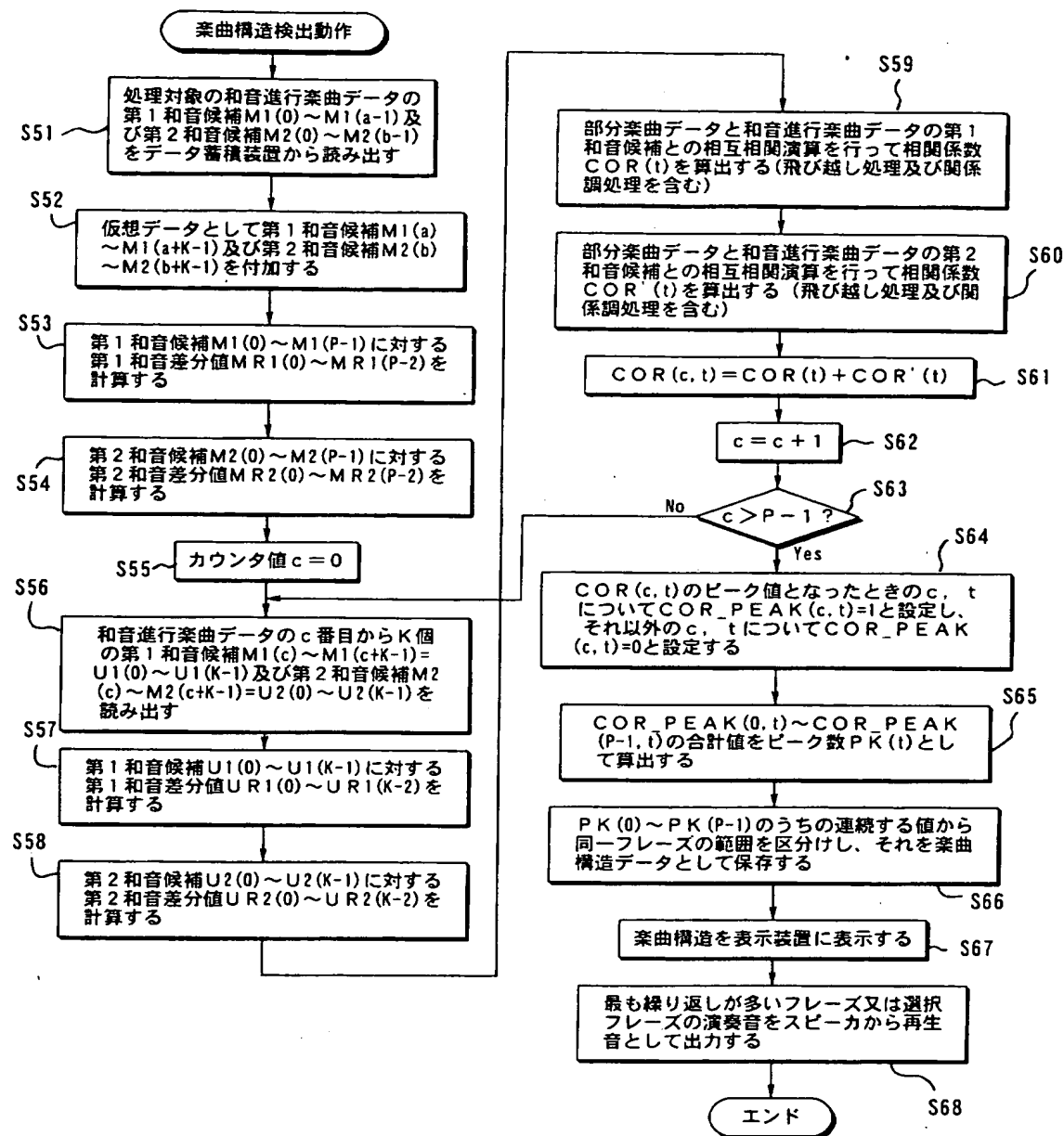
楽曲の識別情報(ファイル名他)(512バイト)

第1和音候補数(4バイト) T1(0) (4バイト)[0x08] T1(1) (4バイト)[0x0A] T1(2) (4バイト)[0x05]
 T1(3) (4バイト)[0x01] T1(4) (4バイト)[0x08].....

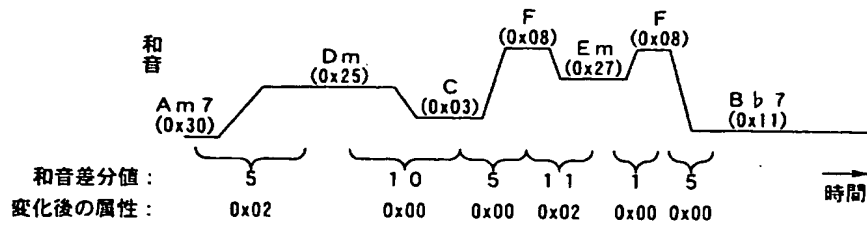
第2和音候補数(4バイト) T2(0) (4バイト)[0x03] T2(1) (4バイト)[0x01] T2(2) (4バイト)[0x29]
 T2(3) (4バイト)[0x01] T2(4) (4バイト)[0x03].....

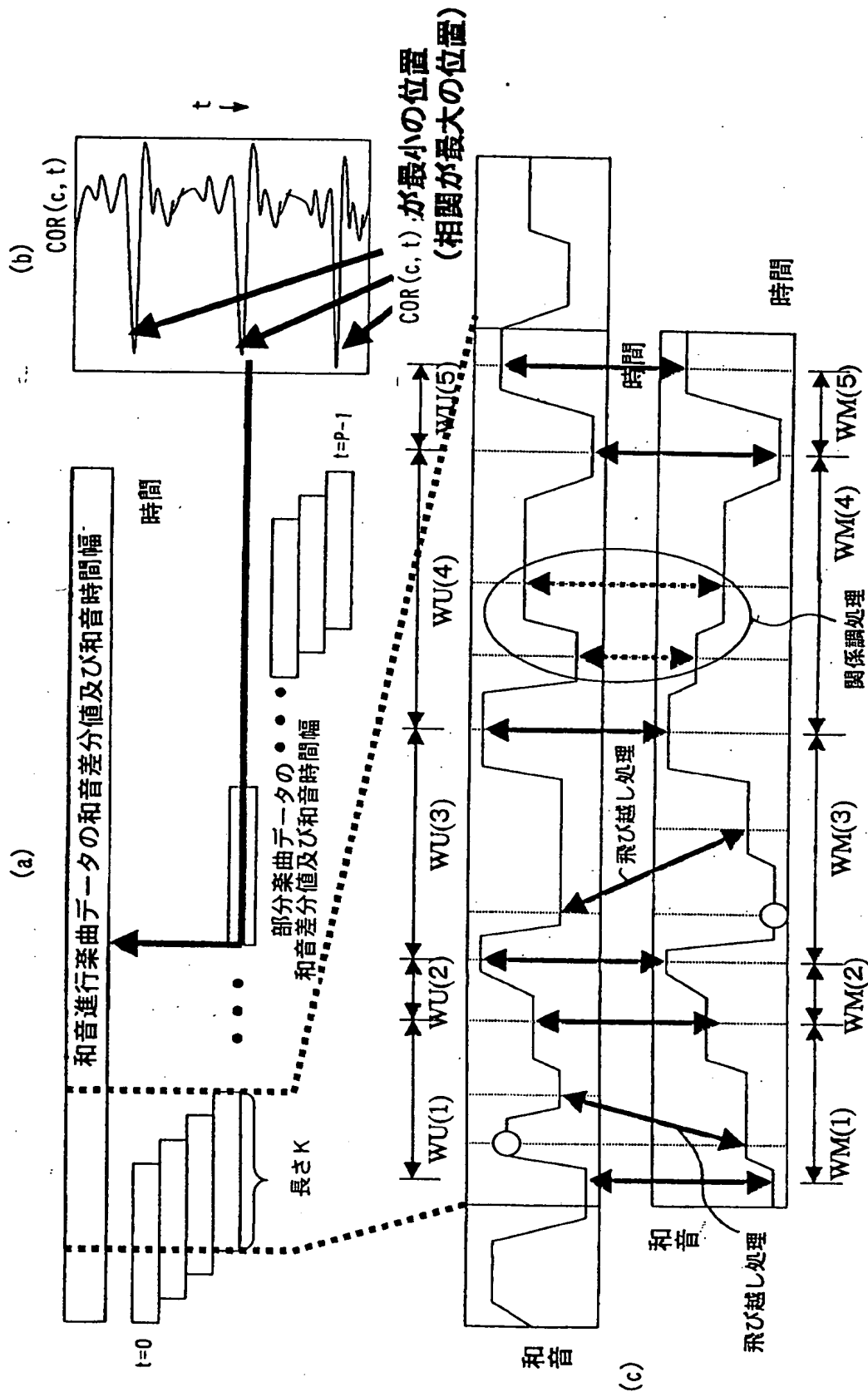
(d)

【図 15】

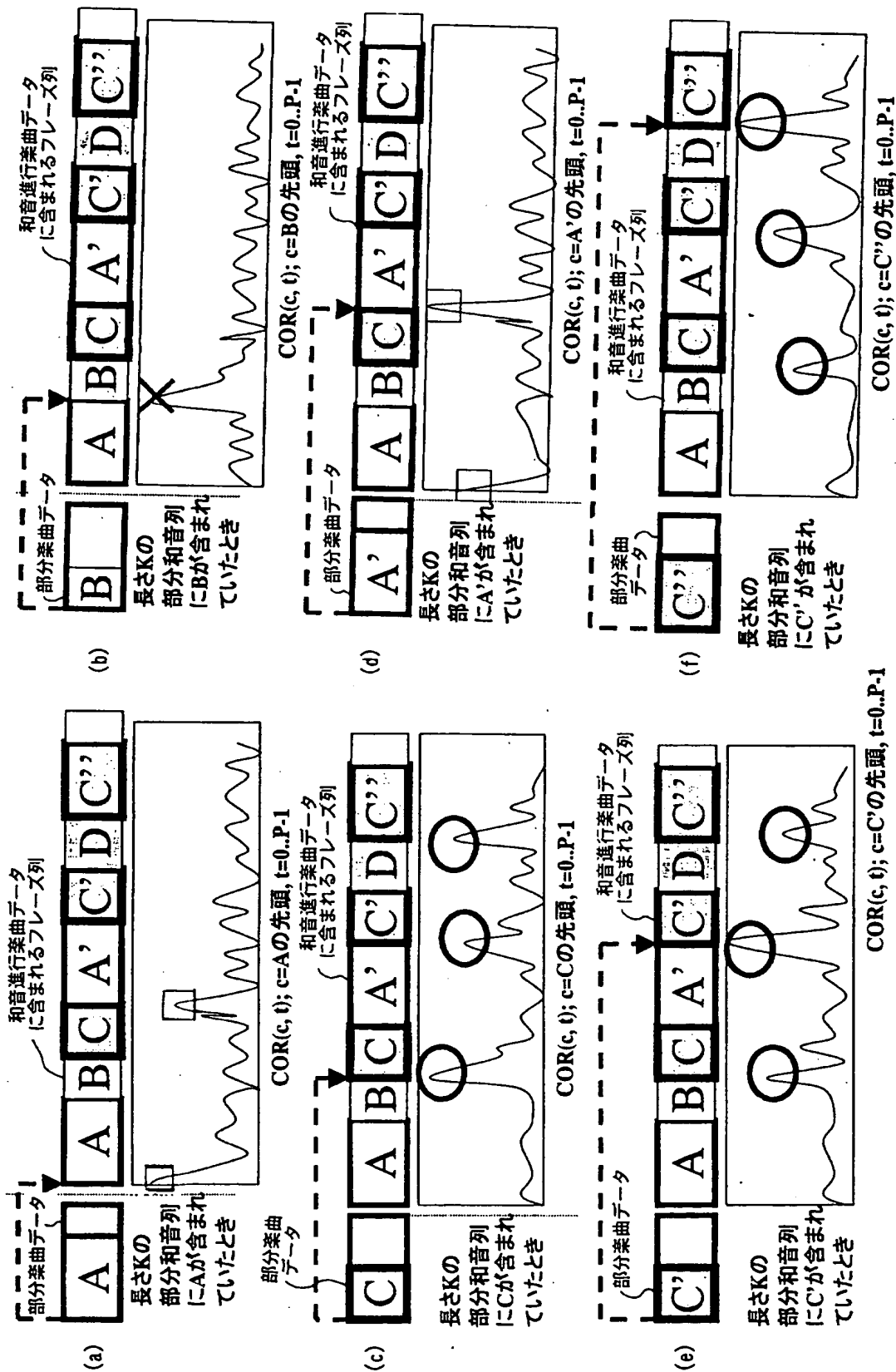


【図 1 6】

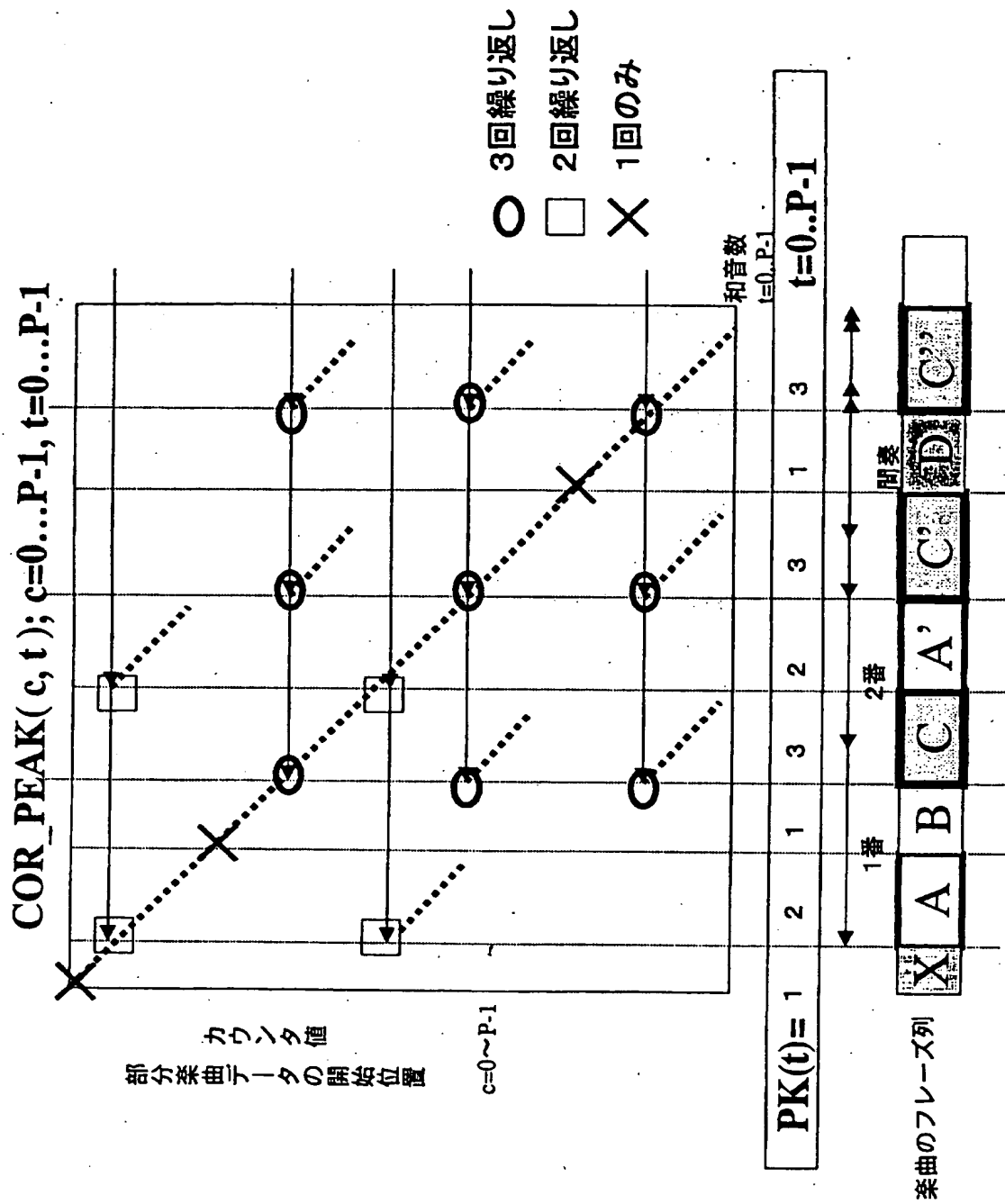




【図 1 9】



【図 2 0】



【図 2 1】

楽曲の識別情報(ファイル名他)(512バイト)

(繰り返し部分1について:3回繰り返されるフレーズ)
 0x03(ランク3、1バイト)第1フレーズ開始時刻情報(4バイト)第1フレーズ終了時刻情報(4バイト)
 第2フレーズ開始時刻情報(4バイト)第2フレーズ終了時刻情報(4バイト)
 第3フレーズ開始時刻情報(4バイト)第3フレーズ終了時刻情報(4バイト)

(繰り返し部分2について:2回繰り返されるフレーズ)
 0x02(ランク2、1バイト)第1フレーズ開始時刻情報(4バイト)第1フレーズ終了時刻情報(4バイト)
 第2フレーズ開始時刻情報(4バイト)第2フレーズ終了時刻情報(4バイト)

(繰り返し部分3について:2回繰り返されるフレーズ)
 0x02(ランク2、1バイト)第1フレーズ開始時刻情報(4バイト)第1フレーズ終了時刻情報(4バイト)
 第2フレーズ開始時刻情報(4バイト)第2フレーズ終了時刻情報(4バイト)

(繰り返し部分4~(実際は繰り返さず、一回のみ出現するもの)について)
 0x01(ランク2、1バイト)フレーズ開始時刻情報(4バイト)フレーズ終了時刻情報(4バイト)
 0x01(ランク2、1バイト)フレーズ開始時刻情報(4バイト)フレーズ終了時刻情報(4バイト)
 0x01(ランク2、1バイト)フレーズ開始時刻情報(4バイト)フレーズ終了時刻情報(4バイト)

.....

【図 2 2】

楽曲タイトルまたはファイル名

聴きどころランキング2

☐ 1発目!

☐ 2発目!

聴きどころランキング1

☐ 1発目!

☐ 2発目!

☐ 3発目!

聴きどころランキング2

1発目! ☐

2発目! ☐

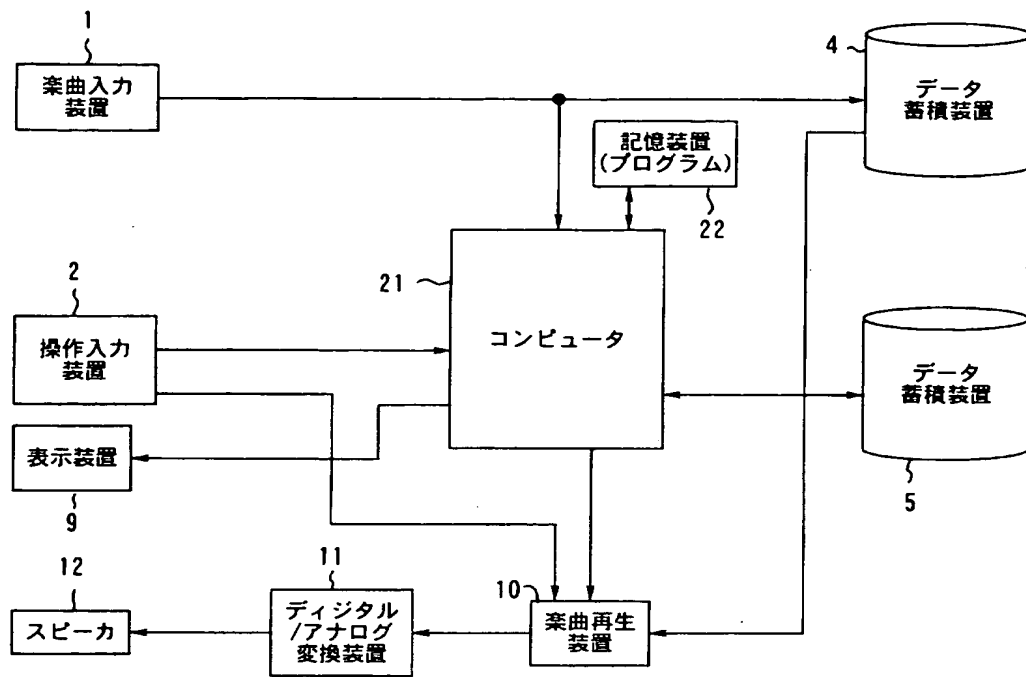
☐ イントロかな?

☐ Bメロかな?

最初から再生!

ストップ!

【図 2 3】



【書類名】 要約書

【要約】

【課題】 繰り返し部分を含む楽曲の構造を簡単な構成で適切に検出することができる楽曲構造検出装置及び方法を提供する。

【解決手段】 楽曲の和音の時系列変化を示す和音進行楽曲データ中の各和音の位置から連続する所定数の和音からなる部分楽曲データを生成し、部分楽曲データ各々と和音進行楽曲データとを和音進行楽曲データ中の各和音の位置から和音変化時の和音の根音変化量と変化後の和音の属性とについて比較して複数の楽曲毎の類似度を算出し、部分楽曲データ毎に算出された類似度各々に応じて類似度が所定値より高いピーク値となった和音進行楽曲データ中の和音の位置を検出し、和音進行楽曲データ中の和音の位置毎に部分楽曲データ全てについて類似度が所定値より高いピーク値となった回数を算出し、その和音の位置毎の算出回数に応じて楽曲構造を示す検出出力を生成する。

【選択図】 図 1

出 願 人 履 歴 情 報

識別番号 [000005016]

1. 変更年月日 1990年 8月31日
[変更理由] 新規登録
住 所 東京都目黒区目黒1丁目4番1号
氏 名 パイオニア株式会社